Research in Autism Spectrum Disorders 32 (2016) 24–42

Contents lists available at ScienceDirect

Research in Autism Spectrum Disorders

journal homepage: http://ees.elsevier.com/RASD/default.asp

Are emotion recognition abilities related to everyday social functioning in ASD? A meta-analysis

Dominic A. Trevisan*, Elina Birmingham

Faculty of Education, Simon Fraser University, 8888 University Dr., Burnaby, BC V5A 1S6, Canada

ARTICLE INFO

Article history:
Received 25 February 2016
Received in revised form 16 August 2016
Accepted 25 August 2016
Number of reviews 3
Available online xxx

Keywords:
Autism
Emotion recognition
Meta-analysis
Social functioning

ABSTRACT

Background: Most developmental theories of autism spectrum disorders (ASD) emphasize a link between the ability to infer others' emotional states with their everyday social functioning. However, rarely has this association been empirically examined in this population.
Methods: We conducted a meta-analysis to quantitatively summarize correlations between performance on facial emotion recognition tasks and theoretically related variables broadly related to social functioning and other cognitive abilities.
Results: Sixty-two correlation coefficients from 27 separate articles met our inclusion criteria. Correlations between the ability to recognize facial expressions (FER) and each category of variables were moderate but significant in the expected direction. FER was positively correlated with age, nonverbal and verbal intelligence, Theory of Mind, and adaptive functioning, and negatively correlated with alexithymia and higher ASD symptoms.
Conclusions: The findings of this meta-analysis indicate that FER abilities represent an important social cognitive ability given its relation to real-world social behavior and other characteristics and cognitive abilities. However, the striking lack of studies in this area calls for more research to gain a clearer understanding of the developmental significance of FER, especially in relation to the broader social impairment characteristic of ASD.

Contents

* Corresponding author.
E-mail addresses: dtrevisa@sfu.ca (D.A. Trevisan), ebirming@sfu.ca (E. Birmingham).

http://dx.doi.org/10.1016/j.rasd.2016.08.004

1. Introduction

1.1. Overview

"Human faces convey information about identity, lip movements, gaze direction, and emotion, and are the primary and most powerful source of information mediating emotional and linguistic communication as well as social interactions" (Gepner, Deruelle, & Grynfeltt, 2001, p. 37).

Autism spectrum disorder (ASD) is a complex neurodevelopmental disorder characterized in part by difficulty forming relationships, and deficits in verbal and nonverbal communication (APA, 2013). In the last three decades an abundance of experimental research has been published that compared the ability of participants with and without ASD to accurately identify the emotional states of others based on their facial expressions, body language and other contextual information. Although findings have been equivocal and dependent on a variety of variables including demographic factors, nature of the task demands, and how the dependent variables are measured (Harms, Martin, & Wallace, 2010), the majority of studies show that the ability to accurately recognize others' emotions is impaired in participants with ASD compared to neurotypical comparison groups (see Gaigg, 2012; Harms et al., 2010; Uljarevic & Hamilton, 2013 for reviews).

It has long been argued that the ability (or lack thereof) to recognize others' emotional states from their facial expressions has wide reaching influences on other areas of social functioning. It is thought that emotion recognition abilities may be a critical mechanism by which observers respond to others empathically and competently, and modify their own behavior adaptively based on the emotional signals of others (Izard et al., 2001). As Jones et al. (2011) described, "Basic emotion recognition is a fundamental building block of more sophisticated emotional and social understanding and establishing the degree of deficit in ASD is important for ascertaining at what level social-emotional understanding begins to break down for these individuals" (p. 275). Despite these purported links between emotion recognition and social functioning more broadly, surprisingly little research has explored which, and to what degree, other variables are continuously related to emotion recognition abilities in individuals with ASD. Instead, the primary motivation of the bulk of research in this area has been to uncover group differences between ASD and control groups, making it difficult to understand the real world impact of emotion recognition difficulties. The ability to decode others' emotional states from nonverbal information such as facial expressions is arguably an important social cognitive ability; yet, it is relatively unknown how this ability contributes to real world social functioning in ASD.

1.2. Autism theories

Many prominent developmental theories of social development in ASD emphasize a closely linked interplay between emotion recognition difficulties and everyday social functioning (Gaigg, 2012). *Mentalizing* accounts (e.g., Baron-Cohen, 1994, 2005; Frith, 2003), for example, suggest that deficient cognitive mechanisms responsible for mentalizing (i.e., inferring the mental states of others, such as beliefs, desires, feelings and intentions) are the underlying cause of the broader social impairment characteristic of ASD. That is, in lacking the ability to understand the thoughts, perspectives and feelings of others, individuals with ASD would find it difficult to respond to others appropriately and empathetically, and to form meaningful connections (Baron-Cohen, Leslie, & Frith, 1985).

The *social motivation* hypothesis (see Chevallier, Kohls, Troiani, Brodkin, & Schultz, 2012, for a review), suggests that social stimuli have lower intrinsic reward value to people with ASD, that diminished social interest is thought to be the underlying cause of impaired social cognitive abilities (such as emotion recognition and theory of mind) and that together these impairments have broad reaching implications for everyday social functioning. The theory holds that experiencing fewer opportunities to interact with social information during critical developmental periods, along with a generally diminished concern or interest in the mental states of others, causes downstream effects on the ability to recognize emotions and other mental states in others (Chevallier et al., 2012). In support of this theory, experimental evidence indicates that individuals with ASD show preference toward non-social objects over people, faces, and eyes while viewing images or video clips (Klin, Jones, Schultz, Volkmar, & Cohen, 2002; Riby & Hancock, 2008), and also show reduced interest toward speech but not non-social sounds relative to control participants (Čeponienė et al., 2003; Klin, 1991). In addition, adults with ASD tend to report less enjoyment and interest in friendships (Baron-Cohen & Wheelwright, 2003), supporting the idea of reduced social motivation. Like *mentalizing* accounts, social motivation accounts make the inherent assumption that social cognitive deficits impact the everyday social functioning of individuals with ASD (although here, these deficits are a *consequence* of diminished social interest, rather than a cause).

The *behavioral self-regulation account* emphasizes that understanding the mental states of others is only beneficial when that information can be used to subsequently regulate one's own social behavior (Bachevalier & Loveland, 2006; Loveland, 2001). While Bachevalier and Loveland's account accords with the common view that failure to perceive others' emotional and mental states may lead to socially incompetent behavior, they argue that social cognitive abilities are not necessarily of primary importance. Rather, lacking the competence to *respond* to others' emotional states in socially appropriate and empathetic ways makes emotional information of little use or interest to people with ASD. Evidence in support of this theory comes from research indicating that some people with ASD do not change their own behavior in response to others' distress, even when they appear to be aware of it (Loveland & Tunali, 1991; Sigman, Kasari, Kwon, & Yirmiya, 1992), and that individuals with ASD may not actively seek out emotional information in their daily lives, even if they can interpret emotional states of others when prompted to do so (Rieffe, Terwogt, & Stockmann, 2000). Thus, the behavioral self-regulation account suggests that regulating one's own social behavior in response to environmental conditions (e.g., responding to the emotional distress of others) is dependent on the evaluation of the functional significance of those conditions for the self, and on the ability to detect that social information in the first place.

Putting their differences aside, these theories share a clear commonality. That is, the difficulties that many individuals with ASD experience with recognizing others' mental states (particularly, emotional states) are thought to contribute to the broader impairments in social functioning that are characteristic of ASD. To provide support for this assumption, it would be expected that performance on emotion recognition tasks would be correlated with measures of everyday social functioning and other theoretically related variables in the ASD population.

1.3. Empirical research

1.3.1. Everyday social functioning

An initial review of the literature suggested very little empirical research has been conducted to provide support for the relationship between emotion recognition abilities and everyday social functioning in the ASD population, and that the sporadic research that has been done has yielded dissimilar results. For example, while some researchers (e.g., Braverman, Fein, Lucci, & Waterhouse, 1989) have reported relationships between emotion recognition and scores on subscales of the Vineland Adaptive Behavior Scales (VABS; Sparrow, Balla, & Cicchetti, 1984) – a measure that largely measures social and communication skills in everyday environments – other have failed to find such associations (e.g., Williams & Gray 2013a). Everyday social functioning is also captured by measures of ASD symptomology, and several studies have examined associations between FER and measures such as the Child Autism Rating Scale (CARS; Schopler, Reichler, de Vellis, & Daly, 1980) and other gold star diagnostic measures. Again, however, the existing data is equivocal; for example, Ozonoff, Rogers, and Pennington (1991), found a nonsignificant correlation between FER and the CARS ($r = -0.25$) in contrast to Tardif, Lainé, Rodriguez, and Gepner (2007) who found a strong, significant association between the same variables ($r = -0.77$).

1.3.2. Theoretically related variables

A handful of other studies have examined relationships between emotion recognition abilities and variables that do not tap into social behavior per se, but have interesting theoretical associations, such as *alexithymia* (e.g., Kätsyri, Saalasti, Tiippana, von Wendt, & Sams, 2008; Ketelaars, Mol, Swaab, & van Rijn, 2016), and *Theory of Mind* (ToM; e.g., Heerey, Keltner, &

Capps, 2003; Rice, Wall, Fogel, & Shic, 2015). Alexithymia, characterized by difficulties interpreting and describing one's own emotions (Nemiah, Freyberger, & Sifneos, 1976), is gaining significant interest among autism researchers (Bird & Cook, 2013), based in part because of evidence suggesting that approximately 50% of individuals with ASD possess high levels of the alexithymia trait compared to 10% in the general population (Berthoz & Hill, 2005). Moreover, a large number of studies have found that higher alexithymia is associated with poorer emotion recognition abilities in the neurotypical population (see Grynberg et al., 2012 for a review). A relationship between impairments in recognizing other's emotions and impairments in recognizing one's own emotions (alexithymia) in ASD would be consistent with the pattern observed by Williams' (2010) that *mentalizing* impairments in individuals with ASD are not isolated to difficulties interpreting *others*' mental states, but also extend to interpreting one's *own* mental states. Similarly, a theoretical relationship exists between emotion recognition and ToM (the understanding that others' beliefs, mental states and perspectives may be different than one's own), as both involve inferring the mental states of others; emotion recognition tasks are specific to inferring the emotional states from body and face cues, whereas ToM tasks typically require participants to infer the perspectives, thoughts and other mental states of characters in a story based on the situational context of that story.

1.3.3. Matching Variables

Several studies examined relationships between FER and both age and intelligence (e.g., Humphreys, Minshew, Leonard, & Behrmann, 2007; Ozonoff, Rogers, & Pennington, 1991; Tantam, Monaghan, Nicholson, & Stirling, 2006). We suspect that the frequency of correlations reported between FER and these variables is largely due to the fact that age and intelligence are used for matching procedures in FER studies, and correlations are often reported as secondary interests. In fact, in many of the studies we reviewed, the examination of correlations between emotion recognition abilities and other variables of interest were not the main aims of the studies.

1.4. Previous meta-analyses

Is it the case that individuals with ASD are in fact poorer at emotion recognition? Two previous meta-analyses found that participants with ASD scored significantly lower than matched neurotypical controls on emotion recognition tasks presenting images of facial expressions and other nonverbal body cues (Uljarevic & Hamilton, 2013) or from tasks presenting facial expression stimuli only (Lozier, Vanmeter, & Marsh, 2014). In their moderator analysis, Uljarevic and Hamilton (2013) found that neither age, intelligence, nor task factors (e.g., labeling vs. matching paradigms) moderated the strength of the overall effect size (Uljarevic & Hamilton, 2013). In contrast, Lozier et al. (2014) found that age, but not intelligence, moderated the strength of the group differences between ASD and controls, such that the ASD deficits *increased* with age. In light of these findings, Lozier et al. (2014) suggested that ASD is associated with functional impairment in the neural architecture that supports face-emotion processing, which may lead to downstream developmental consequences.

Although ours will be the first meta-analysis on relationships between FER and everyday social functioning in ASD, there is some evidence that these relationships exist in neurotypical populations. Hall, Andrzejewski, and Yopchick (2009) conducted a meta-analysis from a combined 215 studies (N = 16,848) to examine the extent to which *interpersonal sensitivity* (i.e., an umbrella term mostly measuring judgment of affective states) correlates with psychosocial variables (e.g., personality traits, indicators of mental health, and social competencies) of the perceiver in neurotypical adults. Hall and colleagues concluded that interpersonal sensitivity was positively correlated with empathy, as well as with more desirable personality traits (e.g., conscientiousness, extraversion and openness), and with a variety of social and work-related competencies. However, the correlations between interpersonal sensitivity and social competencies were very small; for *self*-report social-emotional competence ($r = 0.05$, $p < 0.05$), *other*-report social-emotional competence ($r = 0.16$, $p < 0.001$), and for self-report relationship quality ($r = 0.08$, NS). It should be noted that Hall and colleagues excluded any studies that included participants with psychological impairments or clinical diagnoses.

1.5. The present study

The present meta-analysis extends the work done by these previous meta-analyses, differing from Uljarevic and Hamilton (2013) and Lozier et al. (2014) in that we examined *correlates* of emotion recognition instead of group differences, and differing from Hall et al. (2009) in that we focus only on participants with ASD instead of neurotypical participants. To make sense of the existing literature, we conducted a meta-analysis to quantitatively summarize all studies that examined correlates of the ability to infer others' emotional states from facial information in individuals with ASD. Our primary interest was to investigate the relationship between facial emotion recognition abilities and aspects of everyday social functioning (adaptive functioning and ASD symptomology) because this relationship is heavily emphasized in developmental theory. However, our secondary interest was to summarize previously reported relationships between facial emotion recognition ability and other theoretically relevant variables wherever reported in the extant literature. While everyday social functioning has a strong theoretical link with facial emotion recognition (henceforth referred to as FER), exploring other available correlates that are non-social but cognitive/affective in nature will offer a more complete representation of the traits, characteristics and abilities that covary with FER in this population. Thus, in the present study, we did not selectively choose variables of interest; rather we examined all variables that researchers correlated with FER in the extant literature.

We hope to advance knowledge in at least three different ways. First, examining whether or not there is evidence for a linear relationship between FER and variables related to everyday social functioning will test the theoretical link between these two areas, and the aggregated coefficients for this category will provide a "best estimate" of the degree to which these variables are related in the ASD population. Second, understanding and providing validity for any given construct involves an examination of the degree to which that construct is associated with related and unrelated constructs (e.g., Cronbach & Meehl, 1955; Furr & Bacharach, 2013). As such, summarizing all correlates of FER will provide a clearer picture of the developmental significance of this skillset. Third, while causation cannot be determined from linear correlations, the observed relationships will offer an opportunity to speculate causal pathways among these variables as avenues for future research. Our goal for conducting this study is to summarize the existing data on correlates of FER, and thus we do not offer any specific hypotheses.

2. Methodology

2.1. Article retrieval

Prospective articles were retrieved from reference lists of recently published review articles and meta-analyses (e.g., Harms et al., 2010; Uljarevic & Hamilton, 2013), and databases such as PsycINFO, PsycARTICLES, and GoogleScholar. Database searches included combinations of search terms including variations of the term "autism," (e.g., Asperger, autistic, ASD)"emotion recognition," "facial expressions" and "affect recognition." These searches yielded a total of 581 articles and the lead author reviewed each article. Twenty-seven articles met the inclusion criteria described below.

2.2. Inclusion criteria

We only analyzed emotion recognition performance from face stimuli (FER) and excluded studies that required participants to infer emotion from other types of nonverbal communication (e.g., gestures, vocal tone, body language), music, stories and other non-face stimuli. We made this decision because a) the vast majority of emotion recognition research uses stimuli from faces, and b) it is important to keep the independent variable as succinct as possible in meta-analysis. Additionally, we only included studies that used real faces and excluded stimuli that exclusively used cartoon or hand drawn faces (e.g., Prior, Dahlstrom, & Squires, 1990). In rare cases, we accepted dynamic stimuli (e.g., Gepner et al., 2001), but only if there was no other contextual information besides facial expressions in the clips. For example, studies that used tasks such as the "Reading the Mind in the Films" task (Golan, Baron-Cohen, Hill, & Golan, 2006; Golan, Baron-Cohen, & Golan, 2008) were excluded because the clips may have included utterances from the characters or other contextual clues from the movie scene. Again, to keep the independent variable succinct, we only accepted studies that analyzed FER with basic emotions (i.e., happiness, sadness, joy, disgust, anger and fear) and rejected those that examined complex emotions (e.g., frustration, guilt). In rare cases we did include studies that used combinations of basic and complex emotions in their FER measure. The measures researchers used for correlational analyses reported in this meta-analysis are listed in Appendix A.

Articles were accepted for analysis if they met the following four criteria. First, they must have been published in peer-reviewed, English language journals. Second, they must have conducted an experiment that required participants with ASD to infer the emotion being expressed from face stimuli (FER task). Third, they must have reported a correlation coefficient (Pearson's *r*, Spearman's rank, etc.) for associations between performance on a FER task and another variable of interest. Fourth, they must have reported the correlation specifically for participants who were reported to have a diagnosis on the autism spectrum, including autistic disorder, Asperger's disorder, or pervasive developmental disorder-not otherwise specified. We excluded studies where ASD and neurotypical participants were combined into a single group for correlational analyses. This decision was made because in several studies in which correlations were reported separately for ASD and control groups, very different coefficients were observed, suggesting coefficients collapsed across group levels may not accurately reflect the ASD population. Master and doctoral theses and conference presentations were not included.

2.3. Publication bias and missing data

The tendency for journals to neglect publishing null results (i.e., publication bias) is a concern for all meta-analyses. However, we would argue such bias is less of a concern for our meta-analysis because almost all data we extracted were from articles in which the correlations between FER and other constructs were of secondary importance to the main aims of the article. In fact, nearly half of the coefficients included in our analyses were not statistically significant. In cases where correlations between FER and related constructs were reported as "nonsignificant," efforts were made to estimate the correlation. For example, Pearson's r can be estimated when the exact p-value and sample size are reported (Rosenthal, 1991). However, when researchers fail to include Pearson's r and state, "$p > 0.05$" or "the correlation was nonsignificant" it was impossible to estimate the strength of the correlation, and was therefore not included in our analyses. This happened in only four studies (each from a different category), so it is unlikely that missing data will have significantly exaggerated the effect sizes reported in our results.

Still, it is possible that there are existing data that were unpublished due to publication bias. Relatedly, as correlations between FER and related constructs were of secondary importance in many studies, it is possible that some researchers may have "cherry-picked" the reporting of secondary analyses by preferring to report statistically significant correlations and neglecting nonsignificant correlations. To assess for publication bias in our study, we aggregated all effect sizes (to achieve a reasonable number of effect sizes to test for bias) and used Duval and Tweedie's (2000) *Trim and Fill* procedure using a funnel plot. In Fig. 1, the Y-axis represents the standard error, and the X-axis represents the strength of the effect sizes from each study. Biased sampling distributions tend to contain a disproportionate number of studies on the lower right side of the funnel, because effect sizes with small sample sizes (and thus, higher standard error) are more likely to deviate from the mean due to sampling error (Borenstein, Hedges, Higgins, & Rothstein, 2009). The Trim and Fill procedure uses an iterative process to remove the most extreme small studies from the right side of the plot, re-computing the effect size at each iteration until the funnel plot is symmetric around the new, unbiased effect size. In Fig. 1 the bolded dots indicate that our sample of effect sizes is missing an estimated 7 studies due to publication bias, and that when accounting for these studies, the aggregated effect size (the correlation between FER and a combination of all variable categories) goes down from $r = 0.36$ to 0.34. In conclusion, publication bias is present in our data, but appears to have had little impact.

2.4. Coding of articles

All correlation coefficients reported in the studies that met our inclusion criteria were coded into seven separate categories, including adaptive functioning, ASD symptomology, Theory of Mind (ToM), alexithymia, nonverbal and verbal intelligence, and chronological age (see Appendix A for a full list of the methods researchers used to quantify each variable). Each correlation coefficient was coded into one of these seven discrete categories. While our literature review identified additional correlates beyond these seven variables, we made the decision to categorize variables into the meta-analysis only if at least four separate effect sizes were observed for any one variable.

There were no instances in which more than one correlation coefficient from the same sample were included in any given category, meaning all coefficients within each category are statistically independent. However, in some cases researchers reported correlations between FER and two or more constructs that fit into separate categories. As such, each *category* is not necessarily statistically independent from one another.

2.5. Variable categories

2.5.1. Adaptive functioning

Broadly, adaptive functioning is a multidimensional construct that reflects the personal and social abilities of individuals as they interact with their environment. In the literature we reviewed, two measures were used to assess adaptive

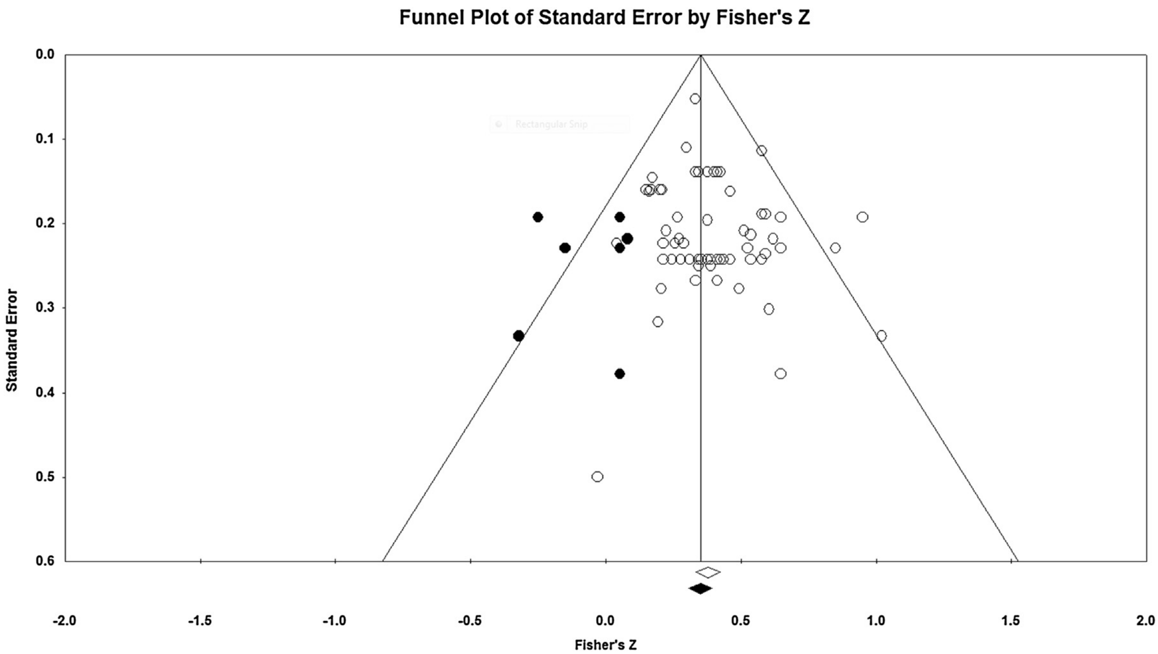

Fig. 1. Funnel plot displaying all present and missing studies. Using the Trim and Fill procedure, the funnel plot indicates the number of unpublished studies missing due to publication bias (indicated in black dots), and the extent to which the overall effect size changes when accounting for the unpublished studies. For this analysis the signs of the ASD Symptomology and Alexithymia coefficients were inverted from negative to positive values to maintain consistency with all other variable categories.

functioning, including the Vineland Adaptive Behavior Scales (Sparrow et al., 1984; Sparrow, Cicchetti, & Balla, 2005) and the Adaptive Behavior Assessment System-II (Harrison & Oakland, 2003). These measures are widely used to measure independent living skills (e.g., eating, dressing, hygiene, personal care, and use of time), aspects of social functioning related to interpersonal behavior (e.g., interpersonal abilities during social interactions, sensitivity to others, responsibility, and use of play and leisure time) and communication (e.g., receptive and expressive communication during face-to-face interactions but all also reading and written communication). The VABS has strong psychometric properties and has been thoroughly standardized for ages across the entire lifespan. Some studies used the full-scale scores of the VABS or ABAS, while others used only the Socialization or Communication subscales. Full-scale scores were used wherever available.

2.5.2. ASD symptomology

This category includes gold standard clinical tools to diagnose and measure symptom severity of autism including the Autism Diagnostic Observation Schedule (ADOS; Lord et al., 2000), Autism Diagnostic Interview – Revised (ADI-R; Lord, Rutter, & Le Couteur, 1994), the Childhood Autism Rating Scale (Schopler et al., 1980) and the Social Responsiveness Scale (SRS, Constantino & Gruber, 2002; Roeyers, Thys, Druart, Schryver, & Schittekatte, 2012). While not typically used for diagnostic purposes, the Autism Quotient (AQ) was also included in this category, as it has shown to reliably discriminate between persons with and without ASD and is commonly used as a screening measure for research purposes (Woodbury-Smith, Robinson, Wheelwright, & Baron-Cohen, 2005). It should be noted that for correlations with FER, some researchers used only subscales specific to social impairment (e.g., the Social and Communication subscales of the ADOS), while other researchers used full-scale scores, which include qualitatively non-social symptoms including repetitive behaviors. For our analyses we combined all studies in this category regardless of whether they used full-scale or subscale scores.

2.5.3. Theory of Mind

Theory of Mind (ToM) here represents a category of performance-based tasks meant to assess participants' ability to infer the mental states, thoughts, intents, desires, feelings, and perspectives of others (Premack & Woodruff, 1978) based on contextual clues of a cartoon or story. Such tasks often expose the participant to more information than what a character in a vignette knows, and participants are required to interpret the perspectives of that character taking into account the differing access to information (e.g., the Smarties False Belief task; Perner, Frith, Leslie, & Leekam, 1989). It should be emphasized that emotion recognition is often considered to be one type of ToM, given that it involves a specific type of mental state inference. For the purposes of our meta-analysis, we make the distinction that FER tasks are entirely *perceptual,* given that participants are afforded no other information besides what is conveyed by facial expressions, whereas tasks in our ToM category typically include vignettes that require higher-order thinking processes such as logic, pragmatic language skills, and the recognition that one's own mind is distinct from another's mind.

2.5.4. Alexithymia

Alexithymia is characterized by a difficulty identifying and describing one's own emotions (Nemiah et al., 1976). Most studies in this category utilized the self-report Toronto Alexithymia Scale (TAS-20; Bagby, Parker, & Taylor, 1994), and one used a subscale from the German translation of the TAS-20's previous version (TAS-26; Kupfer, Brosig, & Brähler, 2001). The TAS-20 is the most commonly used alexithymia measure in the extant literature, and has been extensively validated in neurotypical and clinical adult populations (Taylor, Bagby, & Parker, 1992; Taylor, Bagby, & Parker, 2003). Another study in this category used the Bermond-Vorst Alexithymia Questionnaire (BVAQ; Bermond, 1998).

2.5.5. Verbal intelligence

A variety of measures were used including the Wechsler Intelligence Scale for Children (Wechsler, 1991; Wechsler, 2008), the Stanford–Binet Intelligence Scales (Roid & Barram, 2004), and others. Verbal intelligence measures assess vocabulary, receptive and expressive language, verbal reasoning, and text comprehension, among other related skills (Dyck, Piek, Hay, Smith, & Hallmayer, 2006).

2.5.6. Nonverbal intelligence

In general, subtests from the same measures were used to assess verbal and nonverbal intelligence, respectively. Nonverbal intelligence tests are meant to measure aspects of general cognition that do not require the use of language (DeThorne & Schaefer, 2004), including processing speed, gross motor movements, and perceptual organization abilities such as spatial reasoning and matching tasks.

2.5.7. Age

This category includes only the chronological age of participants.

2.6. Data analysis

Data were analyzed using Comprehensive Meta-analysis Software (CMA; Borenstein, Hedges, Higgins, & Rothstein, 2005). Correlation coefficients were used as our standardized effect for each study. CMA transforms all correlation coefficients into a Fisher's Z value, and then meta-analytically combines these Z values into an overall Z score adjusting for sample size. The resulting Z value is then transformed back into an aggregated correlation coefficient for each category. We used a fixed effects model to combine studies within each category based on the recommendation of Borenstein, Hedges, and Rothstein (2007), that fixed effects models should be used when working with a small number of studies.

Given the dissimilar results observed in the literature reported earlier, it would be useful to identify the variables responsible for the variability in effect sizes using moderator analyses. Specifically, moderator analyses would explore the effects of demographic and methodological variables among individual studies that may account for these differing results. To ascertain whether moderator analyses could be conducted, we examined several dispersion statistics including *Q*-values, I^2 indices, and Tau (*T*) statistics for each category. Observed *Q*-values obtained for all categories were nonsignificant

Study name	Outcome	Statistics for each study				
		Correlation	Lower limit	Upper limit	Z-Value	p-Value
Braverman et al. (1989)	Adaptive Functioning	0.691	0.380	0.862	3.704	0.000
Garcia-Villamisar et al. (2010)	Adaptive Functioning	0.370	-0.101	0.706	1.554	0.120
Wallace et al. (2011)	Adaptive Functioning	0.430	0.141	0.651	2.835	0.005
Williams & Gray (2013a)	Adaptive Functioning	0.206	-0.104	0.480	1.305	0.192
Williams & Gray (2013b)	Adaptive Functioning	0.360	0.105	0.571	2.718	0.007
		0.390	0.253	0.512	5.278	0.000
Williams & Gray (2013a)	Age	0.198	-0.113	0.473	1.253	0.210
Williams & Gray (2013b)	Age	0.320	0.060	0.539	2.392	0.017
Buitelaar et al. (1999) - ASD	Age	0.300	-0.164	0.655	1.276	0.202
Buitelaar et al. (1999) - PDD	Age	0.410	-0.040	0.722	1.796	0.072
Gepner et al. (1996)	Age	-0.030	-0.766	0.740	-0.060	0.952
Gepner et al. (2001)	Age	0.190	-0.403	0.671	0.608	0.543
Katsyri et al. (2008)	Age	0.240	-0.227	0.617	1.009	0.313
Law Smith et al. (2010)	Age	0.530	0.127	0.783	2.504	0.012
Ozonoff et al. (1991)	Age	0.040	-0.378	0.445	0.179	0.858
Rump et al. (2009)	Age	0.330	-0.146	0.682	1.371	0.170
Tanaka et al. (2012)	Age	0.290	0.082	0.474	2.704	0.007
		0.283	0.175	0.385	4.980	0.000
Katsyri et al. (2008)	Alexithymia	-0.210	-0.597	0.256	-0.879	0.379
Cook et al. (2013) - Exp. #1	Alexithymia	-0.457	-0.777	0.050	-1.779	0.075
Ketelaars et al. (2016)	Alexithymia	-0.360	-0.642	0.007	-1.922	0.055
Kliemann et al. (2013)	Alexithymia	-0.265	-0.604	0.155	-1.244	0.213
		-0.320	-0.504	-0.108	-2.911	0.004
Wallace et al. (2011)	ASD Symptomology	-0.159	-0.445	0.156	-0.989	0.323
Williams & Gray (2013a)	ASD Symptomology	-0.166	-0.447	0.145	-1.046	0.295
Williams & Gray (2013b)	ASD Symptomology	-0.330	-0.547	-0.071	-2.472	0.013
Ozonoff et al. (1991)	ASD Symptomology	-0.250	-0.600	0.181	-1.142	0.253
Tanaka et al. (2012)	ASD Symptomology	-0.390	-0.594	-0.139	-2.970	0.003
Cook et al. (2013) - Exp. #1	ASD Symptomology	-0.202	-0.634	0.326	-0.738	0.460
Kliemann et al. (2013)	ASD Symptomology	-0.550	-0.780	-0.188	-2.834	0.005
Bal et al. (2010)	ASD Symptomology	-0.540	-0.832	-0.013	-2.004	0.045
Evers et al. (2015)	ASD Symptomology	-0.170	-0.428	0.114	-1.177	0.239
Humphreys et al. (2007)	ASD Symptomology	-0.520	-0.782	-0.101	-2.376	0.017
Rice et al. (2015)	ASD Symptomology	-0.530	-0.745	-0.216	-3.123	0.002
Tardif et al. (2007)	ASD Symptomology	-0.770	-0.932	-0.351	-3.061	0.002
Grove et al. (2014)	ASD Symptomology	-0.320	-0.409	-0.224	-6.293	0.000
		-0.330	-0.394	-0.263	-9.110	0.000
Braverman et al. (1989)	Nonverbal Intelligence	0.570	0.195	0.799	2.822	0.005
Williams & Gray (2013b)	Nonverbal Intelligence	0.380	0.128	0.586	2.885	0.004
Buitelaar et al. (1999) - ASD	Nonverbal Intelligence	0.390	-0.063	0.710	1.698	0.090
Buitelaar et al. (1999) - PDD	Nonverbal Intelligence	0.240	-0.227	0.617	1.009	0.313
Katsyri et al. (2008)	Nonverbal Intelligence	0.330	-0.132	0.674	1.414	0.158
Ozonoff et al. (1991)	Nonverbal Intelligence	0.210	-0.221	0.573	0.953	0.340
Humphreys et al. (2007)	Nonverbal Intelligence	0.490	0.061	0.766	2.210	0.027
Dyck et al. (2006) - Exp. #2	Nonverbal Intelligence	0.260	-0.111	0.567	1.383	0.167
Tantam et al. (2006)	Nonverbal Intelligence	0.570	-0.093	0.883	1.713	0.087
Teunisse & de Gelder (2001)	Nonverbal Intelligence	0.321	-0.189	0.694	1.245	0.213
Wallace et al. (2008)	Nonverbal Intelligence	0.220	-0.183	0.560	1.073	0.283
		0.354	0.236	0.461	5.604	0.000
Buitelaar et al. (1999) - ASD	Theory of Mind	0.430	-0.015	0.733	1.896	0.058
Buitelaar et al. (1999) - PDD	Theory of Mind	0.370	-0.087	0.698	1.602	0.109
Rice et al. (2015)	Theory of Mind	0.520	0.203	0.738	3.050	0.002
Dyck et al. (2006) - Exp. #2	Theory of Mind	0.740	0.518	0.869	4.939	0.000
Buitelaar & VanderWees (1997)	Theory of Mind	0.400	-0.052	0.716	1.747	0.081
Heerey et al. (2003)	Theory of Mind	0.490	0.118	0.742	2.514	0.012
		0.528	0.392	0.642	6.648	0.000
Braverman et al. (1989)	Verbal Intelligence	0.481	0.074	0.750	2.285	0.022
Williams & Gray (2013a)	Verbal Intelligence	0.149	-0.162	0.433	0.937	0.349
Williams & Gray (2013b)	Verbal Intelligence	0.400	0.151	0.601	3.055	0.002
Buitelaar et al. (1999) - ASD	Verbal Intelligence	0.340	-0.121	0.680	1.460	0.144
Buitelaar et al. (1999) - PDD	Verbal Intelligence	0.270	-0.196	0.636	1.142	0.254
Katsyri et al. (2008)	Verbal Intelligence	0.360	-0.098	0.692	1.554	0.120
Ozonoff et al. (1991)	Verbal Intelligence	0.280	-0.149	0.621	1.287	0.198
Humphreys et al. (2007)	Verbal Intelligence	0.520	0.101	0.782	2.376	0.017
Dyck et al. (2006) - Exp. #2	Verbal Intelligence	0.570	0.264	0.772	3.365	0.001
Teunisse & de Gelder (2001)	Verbal Intelligence	0.390	-0.112	0.733	1.541	0.123
Wallace et al. (2008)	Verbal Intelligence	0.470	0.101	0.725	2.446	0.014
Loveland et al. (2008)	Verbal Intelligence	0.520	0.339	0.664	5.057	0.000
		0.415	0.324	0.499	8.141	0.000

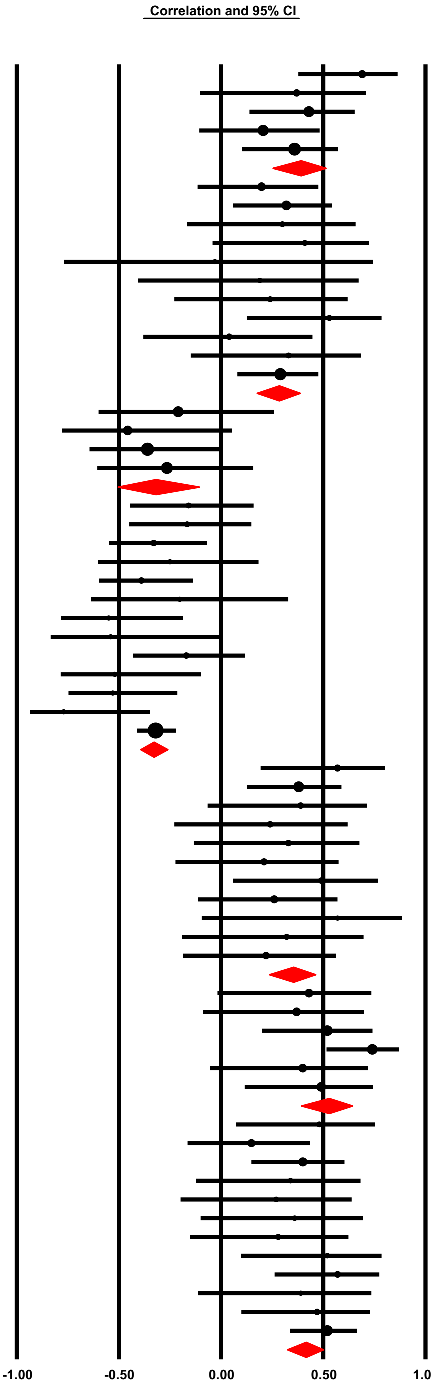

Fig. 2. Forest Plot of all Effects Analyzed in Meta-analysis.

(*ps* > 0.10), indicating an absence of heterogeneity. The I^2 indices, which indicate the proportion of observed variance reflected by variance in *true* effect sizes rather than sampling error (Borenstein et al., 2009), were all close to 0%. Finally, the Tau statistics, which indicate the standard deviation of true effect sizes for all categories, were close to zero (all less than 0.12). Taken altogether, this pattern of results precludes the possibility of running moderator analyses. It should be noted that the actual range of effect sizes (reported in Section 3.1 and observed in Fig. 2) within any given category was often quite large. In light of the discrepancy between the dispersion statistics cited above and the wide range of observed effect sizes, it is impossible to infer from our data whether the heterogeneity in observed effect sizes reflects *true* variance in the effect sizes or if this heterogeneity is due to random sampling error.

3. Results and discussion

In total, 62 correlation coefficients from 28 independent samples reported in 27 different articles met our inclusion criteria. The 62 coefficients were categorized into the seven separate variable categories described earlier. The results are summarized in Table 1. The combined number of participants with ASD across the 28 separate samples was N = and the mean age of the entire sample was 26.3 years. The mean age among individual samples ranged from 5.2 to 41.4 years. A forest plot displaying all coefficients included in this study, along with confidence intervals and *p*-values for each study are displayed in Fig. 2. Fig. 3 shows the sampling distribution of effect sizes across all studies. Appendix A details specific information about each study from which meta-analytic information was extracted.

3.1. Effect sizes and possible explanations

3.1.1. What is the relationship between FER and adaptive functioning?

Five studies were included in this category. The aggregated effect was $r = 0.39$, 95% CI[0.25, 0.51], $p < 0.001$. Of the five studies, the strength of the correlations ranged from $r = 0.21$ to 0.69. These results accord with the assumption of the theories presented earlier, that recognizing others' emotional states may be a critical component of social functioning necessary for communicating effectively and interacting with others. Thus, at first glance, our finding appears to provide some support for this contention. On the other hand, one could speculate a bi-directional relationship between FER and adaptive functioning, such that people with better adaptive functioning find themselves in more social situations, which increases their exposure to a diverse range of facial expressions aiding their development of FER capabilities.

3.1.2. What is the relationship between FER and ASD symptomology?

Thirteen studies were included in this category. The aggregated effect was $r = -0.33$, 95% CI[−0.39, −0.26], $p < 0.001$. Of the thirteen studies, the strength of the correlations ranged from $r = -0.13$ to −0.77. This negative relationship was expected given that *lower* FER abilities are theoretically associated with *higher* levels of ASD symptoms. It should also be noted that this category overlaps somewhat with the adaptive functioning category, given that the diagnostic tools and non-diagnostic measures included in the ASD Symptomology category largely quantify ASD-related social impairment. While Uljarevic and Hamilton's (2013) and Lozier et al.'s (2014) meta-analyses demonstrated emotion recognition deficits in ASD compared to control participants when examining *between-group* comparisons, the findings of our meta-analysis complement their findings, by indicating that ASD symptom severity and FER co-vary *within* the ASD population.

3.1.3. What is the relationship between FER and Theory of Mind abilities?

Six studies were included in this category. The aggregated effect was $r = 0.53$, 95% CI[0.39, 0.64], $p < 0.001$. Of the six studies, the strength of the correlations ranged from $r = 0.37$ to 0.74. The association between FER and ToM is the strongest of the seven categories. This strong relationship is expected given that both FER and ToM tasks are performance-based measures that require the participant to infer the mental states of others, based on a variety of cues, suggesting both may be facilitated by similar cognitive mechanisms. Although we ensured that none of the ToM measures directly assessed FER (all ToM measures involved interpreting vignettes or cartoons, and not facial expressions) some of the ToM tasks may have assessed other aspects of emotional understanding. For instance, in some ToM tasks participants were required to infer

Table 1
Correlations between emotion recognition performance and variable categories.

Category	Pearson's *r*	95% CI Lower Upper	*SE*	Z-value	*k*	*N*
Adaptive Functioning	0.39**	0.25 0.51	0.03	5.28	5	179
ASD Symptomology	−0.33**	−0.39 −0.26	0.01	−9.11	13	746
Theory of Mind	0.53**	0.40 0.64	0.03	6.65	6	146
Alexithymia	−0.32*	−0.50 −0.11	0.04	−2.91	4	89
Verbal Intelligence	0.42**	0.32 0.50	0.02	8.14	12	375
Nonverbal Intelligence	0.35**	0.24 0.46	0.02	5.60	11	263
Age	0.28**	0.18 0.39	0.03	4.98	11	208

Note. SE = Standard error, k = number of studies per category, N = number of participants per category, CI = Confidence interval.
* $p < 0.005$.
** $p < 0.001$.

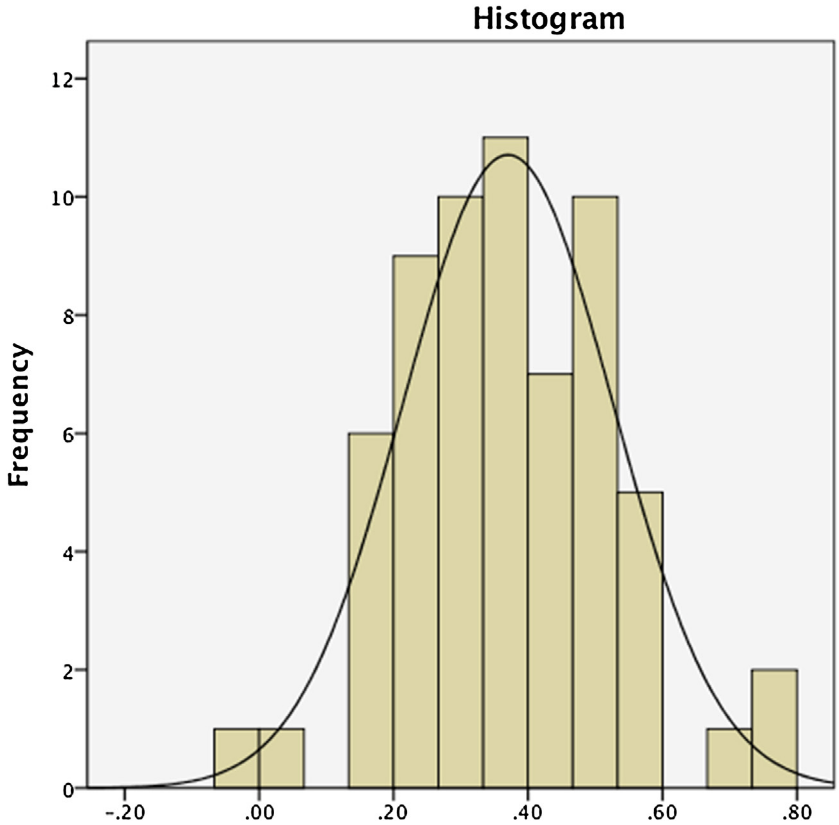

Fig. 3. Distribution of Effect Sizes Across all Categories. For this distribution of effect sizes the signs of the ASD Symptomology and Alexithymia coefficients were inverted from negative to positive values to maintain consistency with the other variable categories.

emotional and other mental state information based on the contexts of a story or vignette. Thus, it is possible that the aggregated effect size for this category is exaggerated due to shared content between the ToM and FER tasks. Again, however, we wish to emphasize that FER tasks were entirely *perceptual*, requiring participants to decode facial expressions with no other information, whereas emotional ToM tasks required participants to use higher-order thinking skills to infer emotional states based on contextual cues in the absence of facial expressions or other nonverbal body cues.

3.1.4. What is the relationship between FER and alexithymia?

Four studies were included in this category. The aggregated effect was $r = -0.32$, 95% CI[−0.50, −0.11], $p = 0.004$ (in the expected direction). Of the four studies, the strength of the correlations ranged from $r = -0.20$ to −0.46. Consistent with our results focusing only on the ASD population, a recent systematic review (Grynberg et al., 2012) found that alexithymia is associated with deficits in emotion recognition abilities across an array of clinical and healthy populations. The reason for this relationship is unclear although it is possible that this association is explained by a domain-general deficit in emotional state awareness, not necessarily restricted to others. In other words, individuals who have difficulties identifying emotions in others (i.e., *low* FER) may also be likely to experience difficulties understanding emotions in themselves (i.e., *high* alexithymia). However, this explanation is insufficient, as the sources of information needed to identify one's own emotions are quite different than the sources needed to identify others' emotions. FER, for example, requires the observer to arrive at a mental state attribution through decoding visual perception information from the face, whereas inferring one's own emotional states requires introspection into one's own thoughts and physiological indicators (e.g., increased heart rate for anger). Future research is needed in this area to gain a more sophisticated understanding of the relationship between *self* and *others*' emotional state attributions.

3.1.5. What is the relationship between FER and verbal intelligence?

Twelve studies were included in this category. The aggregated effect was $r = 0.42$, 95% CI[0.32, 0.50], $p < 0.001$. Of the twelve studies, the strength of the correlations ranged from $r = 0.15$ to 0.57. This relationship is perhaps not surprising (Harms et al., 2010), as many FER paradigms require some verbal ability, especially emotion-word vocabulary. *Labeling* paradigms, for example, require participants to choose the correct emotion label, either from a set list or, as in the case of free-labeling tasks, from memory. Thus, those with stronger verbal ability would be better able to (a) access stored labels for identified facial expressions; and/or (b) use their verbal strengths to infer the correct answer from reading the choices, even if the expressed emotion was not immediately recognized.

3.1.6. What is the relationship between FER and nonverbal intelligence?

Eleven studies were included in this category. The aggregated effect was $r = 0.36$, 95% CI[0.24, 0.46], $p < 0.001$. Of the eleven studies, the strength of the correlations ranged from $r = 0.21$ to 0.57. It is possible that FER performance is dependent in part on spatial reasoning skills, a major component of most nonverbal intelligence batteries. Facial expressions of even the most basic emotions are interpreted by complex combinations and positioning of a large number of critical “action units” within the face (see Ekman, Friesen, & Hager, 2002), so accurate interpretation of facial expressions may be dependent in part on subconscious processing of how each of these action units interact to represent various emotional states.

3.1.7. What is the relationship between FER and chronological age?

Eleven studies were included in this category. The aggregated effect was $r = 0.28$, 95% CI[0.18, 0.39], $p < 0.001$. Of the eleven studies, the correlations ranged from $r = -0.03$ to 0.53. This positive association was expected as FER abilities presumably develop throughout the lifespan as a result of maturation and accumulating practice in real-world social interactions (e.g., Birmingham et al., 2013).

3.2. Differences in effect sizes among categories

All constructs are embedded in a theoretical context, and the conceptual foundation of any given construct is intricately connected with the degree to which that construct is associated with other related constructs (Furr & Bacharach, 2013). Therefore, it is useful to consider the relative strength of associations between FER and each variable category in comparison to one another to gain a clearer picture of what the FER construct comprises. It also logically follows that the associations between FER and similar measures (convergent validity) should be more strongly correlated than between FER tests and dissimilar measures (discriminant validity). This can be statistically examined within our meta-analysis by comparing the aggregated effect sizes of any two categories along with their 95% confidence intervals (all negative signs reversed for consistency). If the confidence intervals of the effect sizes do not overlap, the strength of the correlations can be considered significantly different. For example, ToM had the strongest association with FER compared to any of the other variables, $r = 0.53$, CI[0.40, 0.64], and the strength of this effect size is significantly stronger than the effect sizes for age and ASD symptomology because the *lower* limit of the ToM CI (0.40) is higher than the *upper* limit of the CIs for each of these variables (both = 0.39). Because ToM arguably has the strongest theoretical overlap with FER (as both require accurate interpretation of others' mental/emotional states), the fact that the association between these two variables is the strongest of all the other categories provides support for the convergent validity of the FER construct. No other effect size comparisons were statistically significant.

3.3. Method variance

A final consideration to note is the psychometric concept of *shared method variance.* Regardless of what constructs are being correlated, correlations between two variables tend to be stronger when they are measured using similar or identical methods, and tend to be weaker when they were measured using dissimilar methods (Furr & Bacharach, 2013). Examples of methods used to measure the variables reviewed in this meta-analysis include performance-based, self-report and other-report (e.g., parent, teacher or clinician ratings) measures. Because FER is a performance-based measure, one might expect to see the strongest correlations between it and the other performance-based variables, including ToM, verbal intelligence and nonverbal intelligence, because all such variables, to some degree, capture *test-taking aptitude.* Indeed, of the seven variable categories, ToM, verbal intelligence, and nonverbal intelligence yielded the first, second and fourth largest effect sizes, respectively. All of the other variable categories (except for chronological age) were measured mostly, if not entirely, using self-report or other-report methods. Although it is impossible to decipher from our data how much of the correlation strengths between FER and each, ToM, verbal, and nonverbal intelligence reflect true associations between the underlying constructs and how much is reflected by shared method variance, researchers, and consumers of research, should consider the role of method variance when interpreting correlation strength.

3.4. General discussion

The developmental theories of ASD we described in the introduction strongly emphasize the role of emotion recognition deficits in contributing to difficulties in everyday social functioning. The main rationale for this study was to test this theoretical assumption by summarizing previously reported correlates of performance on facial emotion recognition tasks (FER) in individuals with ASD using meta-analysis. The correlates that we examined in relation to FER included aspects of everyday social functioning, and other theoretically relevant variables. Of the seven categories of correlates we summarized in Table 1, two in particular reflected important aspects of everyday social functioning: “adaptive functioning” and “ASD symptomology.” Importantly, these two categories measure aspects of *behavior* while the other variable categories (aside from *age*) reviewed in this meta-analysis represent either cognitive or affective processes. The measures of adaptive functioning included content related to "Daily Living Skills" (e.g., eating, dressing, hygiene, personal care, and use of time), "Communication" (e.g., receptive and expressive communication during face-to-face interactions, as well as reading and written communication), and "Socialization" (e.g., interpersonal abilities during social interactions, sensitivity to others, responsibility, and use of play and leisure time). Evidently, much of the content of the VABS assesses aspects of personal and social functioning that are relevant to the real-world challenges individuals with ASD experience (Volkmar et al., 1987). Relatedly, the measures of ASD symptomology we examined included social behavior deficits specific to ASD such as atypical joint attention, reciprocity, empathetic gestures, use of facial expressions to facilitate social interaction, eye contact, mannerisms, speech, and conversational turn-taking. Thus, in addition to real-life indices of social "outcomes" as assessed by the adaptive functioning measures, FER appears to be correlated with the finer-grained behavioral nuances of social interactions assessed by measures of ASD symptomology.

That FER appears to be moderately related to important indicators of everyday social functioning among a large number of individuals with ASD reviewed in this meta-analysis provides some support for the assumption of the theories reviewed in the Introduction. Although causal mechanisms cannot be inferred from simple correlations, the findings of this meta-analysis suggest FER is associated with measures of everyday social functioning, and provide indirect support for the idea that impaired FER in ASD may be contributing to broader social impairment and interpersonal abnormalities that are characteristic of ASD. Admittedly, the strength of the relationships between FER and everyday social functioning are moderate at best ($r = 0.39$ for adaptive functioning, $r = -0.33$ for ASD symptomology). However, this is not surprising when considering other aspects of social interactions besides decoding emotional information that are necessary for successful social interactions. Nowicki and Duke (2013) offer a useful conceptual model describing how nonverbal emotional information is used during social interactions. First, an individual needs to be able to *discriminate*, or recognize that subtle variations in facial expression can represent entirely different emotions. Second, an individual must be able to *decode* emotional information, (e.g., by understanding that folded arms, furrowed eyebrows and tight lips may indicate anger). Third, an individual must have the ability to make sense of that emotional information within the situational context of the social interaction (e.g., understanding why another person is angry, and to whom that anger is directed). Finally, the individual must *apply* the information learned in the previous steps by modifying one's own behavior and nonverbal communication according to the specific demands of the situational context. We would care to add to this model, preceding step 1, *attention* to relevant nonverbal cues, which has shown to be reduced in ASD (Rieffe et al., 2000; see also Birmingham, Ristic, & Kingstone, 2012). Clearly, the ability to decode others' facial expressions (i.e., FER) represents only one component of a larger process necessary for *using* that information in "socially competent" ways. A challenging, but useful area for future research will be to systematically examine each of these steps separately to gain a clearer understanding of the specific task demands of ongoing social interactions that may break down in individuals with ASD. Results from our meta-analysis seem to suggest that FER is one critical aspect of this larger process.

The majority of our variable categories (ToM, alexithymia, verbal IQ and nonverbal IQ) were cognitive or affective in nature, as opposed to measuring real-world behaviors. The observed association between FER and alexithymia offers exciting new areas for future research. Williams (2010) has commented that individuals with ASD may have a generalized *mentalizing* impairment that extends to both the self and others, which appears to parallel our finding that understanding emotions may also extend to both the self and others. A related possibility for explaining the relationship between FER and alexythmia, is that when observers attempt to identify others' emotional states, their judgments are aided by replicating or *simulating* those emotions in themselves (Goldman & Sripada, 2005). An interesting possibility for future research is whether alexithymia disrupts the route by which simulation occurs, thus explaining comorbid deficits in FER and other mental state attributions.

Both verbal and nonverbal IQ were moderately associated with FER, which highlights the importance of matching participant groups or otherwise controlling for IQ in FER studies. In fact, verbal IQ had the second strongest association with FER ($r = 0.42$) of all the variable categories, only behind ToM, suggesting individuals with ASD who have proficient verbal abilities may be applying their cognitive and linguistic resources to aid performance on FER tasks (Harms et al., 2010). The last category, not related to social functioning, cognitive or affective processes, was *age*. We found that FER was associated with age, but the aggregated effect size was the smallest of all variable categories ($r = 0.28$). Lozier et al. (2014) found that group differences in FER between ASD and neurotypical controls *increased* as a function of age. Taken together with our findings, it appears that FER improves in individuals with ASD as a function of age, but at a lesser rate than neurotypical individuals.

3.5. Clinical implications

Can FER abilities be taught to people with ASD, and if so, are these initiatives worthwhile? Computerized programs with games and assessment tools designed for this purpose are becoming increasingly common (e.g., Golan & Baron-Cohen, 2006; Golan et al., 2010; Kandalaft, Didehbani, Krawczyk, & Chapman, 2013; Silver & Oakes, 2001), and most of them show modest gains in FER performance in their participants with ASD. An extensive examination of the effectiveness of these interventions is beyond the scope of this article, but according to Reichow and Volkmar's (2010) review, most of these interventions do not meet their criteria for "Evidence-based practices" (cf; Golan & Baron-Cohen, 2006) – the most common criticisms relating to deficient evidence of generalizability and/or long- term maintenance of treatment gains. The results of our meta-analysis provide important motivation for rigorous research on such interventions, because our data suggest FER abilities may impact social functioning. There is an urgent need for future research to identify whether gains in FER as a result of intervention can be generalized to positively impact real-world social outcomes.

3.6. Limitations and future research

There are several limitations of the present study that should be addressed. The aggregate effect sizes reported in our study were from a relatively small number of studies in each category (ranging from 4 to 13), so it is difficult to state with confidence whether the effect sizes reported in this study are generalizable to the entire ASD population. However, meta-analyses can be conducted with as few as two studies (Valentine, Pigott, & Rothstein, 2010), and should be published at varying levels of maturity for any given research area to summarize the existing literature and to guide future research. Similarly, the sample sizes of individual studies were often low; yielding correlation coefficients that may be especially susceptible to measurement error. However, one of the major benefits of meta-analysis is that it tends to cancel out the effects of measurement error in individual studies to provide aggregate effect sizes that may more accurately reflect the population (Hunter & Schmidt, 2004). In light of the common practice of utilizing small sample sizes in autism research, meta-analysis will continue to be an important method for summarizing empirical research in this area.

Another limitation is that we could not conduct moderator analyses, as effect size heterogeneity was attributed to measurement error. For example, it would have been interesting to examine whether variables such as intelligence and age moderated the strength of correlations between FER and theoretically related variables, such as social competence, ToM, and empathy. This remains an open question for future research. Another important question unanswered from our study, is whether methodological features (e.g., labeling versus matching paradigms and dynamic versus static stimuli) moderate the relationship between FER and other variables, which would yield important information regarding the ecological validity of various FER paradigms. For example, we could not test Harms et al.'s (2010) suggestion that labeling paradigms may be more strongly influenced by verbal intelligence (compared to matching paradigms). Additionally, as most studies collapsed FER performance across all individual emotions before conducting correlational analyses, we could not examine whether there are certain emotions that are more or less strongly associated with everyday social functioning. For instance, Wallace et al. (2011) found that recognition of sadness, but none of the other basic emotions (happiness, anger, fear, disgust or surprise) was uniquely related to adaptive functioning and ASD symptomology, whereas Williams and Gray (2013b) found that fear and anger, but not happiness and sadness, were associated with ASD symptomology. Future research should explore these possibilities in greater detail.

Finally, many of the measures included in our variable categories (e.g., ASD symptomology and ToM measures) assessed aspects of emotion processing. For our analyses it would have been impossible to separate emotional and non-emotional items on these measures, but it is possible that different facets of emotional understanding (e.g., FER, emotion regulation, nonverbal emotional expression) are not entirely distinct constructs, and therefore may conflate correlations between FER and some of the variables included in this meta-analysis. Future research could examine this possibility, for example by separating emotional and non-emotional items of ASD symptomology measures and running correlations separately.

3.7. Methodological considerations

Here we offer researchers several methodological considerations. First, almost all published studies that examined correlates of FER conducted simple linear correlations, precluding the possibility of examining causal pathways. To gain a more sophisticated conceptualization of the developmental significance of emotion recognition abilities in the ASD population, future research should utilize longitudinal research designs, structural equation modeling techniques, and consider the roles of confounding and mediating variables. Second, researchers should shift towards a "meta-analytic thinking" orientation (Cumming, 2013; Henson, 2006), guided by the principle that results of all statistical tests should be reported along with effect sizes and confidence intervals, with less emphasis on null hypothesis significance testing. Reporting effect sizes (including nonsignificant correlation coefficients) allows the possibility of including all relevant works into meta-analyses, and decreases the potential influence of publication bias. This is especially important for research with special populations such as ASD, as sample sizes are often quite low. A small or moderate correlation, for instance, may be *practically* or *clinically* significant even if it is not *statistically* significant. Third, in the studies we reviewed, researchers commonly created their own measurement tools or manipulated previously validated tools for their own specific research interests. Researchers should report the internal consistencies of the measures they use in their own studies and carefully consider how low reliabilities of measurement tools may affect their results. In almost all cases, low reliabilities will attenuate correlation strength (Furr & Bacharach, 2013). It is possible that low reliabilities contributed to excessive measurement error in the present meta-analysis, although it was not possible to examine this possibility in depth as only four of the studies included in our meta-analysis reported reliabilities of the FER tools they used. Finally, as FER data were collected using static images in almost all studies we reviewed, researchers should develop innovative strategies to explore FER in more ecologically valid contexts in order to gain a more sophisticated understanding of the specific ways in which FER deficits may contribute to the social impairments characteristic of ASD.

4. Conclusion

The purpose of this meta-analysis was to quantitatively summarize previously published correlations between FER and measures of everyday social functioning, as well as with other theoretically related variables in participants with ASD. All associations reviewed were statistically significant in the expected direction. Our findings indicate that FER is on average, positively associated with a variety of desirable qualities including intelligence, adaptive functioning, and the ability to infer others' mental states (ToM), and negatively associated with problematic qualities including difficulties understanding emotions in oneself (alexithymia) and higher levels of ASD symptoms.

The results of our meta-analysis offer some support for the validity of the FER construct in participants with ASD, as all correlations summarized in Table 1 were significant in the expected direction and consistent with theory. However, our central interest was to test the relationships between emotion recognition abilities and aspects of everyday social functioning emphasized in autism theory, and we were surprised by the relatively small number of studies that had attempted to empirically examine this theoretical assumption. This gap in the literature is likely explained by the notable lack of measures used to assess social functioning (Yager & Iarocci, 2013). The relative paucity of research examining associations with FER is concerning, and more research is needed to better understand the developmental significance of this skillset.

Acknowledgements

We thank Maureen Hoskyn for her helpful comments on early drafts of our paper. This work was supported by a grant (639947) to EB from the Social Sciences and Humanities Research Council of Canada (SSHRC).

Appendix A.

Summary of Studies that Examined Correlates of FER in ASD Participants.

Study	N (males)	Diagnosis	Mean age (years)	FER Type	FER Stimuli	Emotions measured	Variable Category	Measure	Pearson's *r* estimate (p)
Braverman et al. (1989)	22 (NR)	ASD/PDD	10.8	EM	EFAS	Basic Emotions	Adaptive Functioning	VABS	0.69 (<0.001)
García-Villamisar, Rojahn, Zaja, and Jodra, (2010)	19 (11)	ASD	32.4	EL/ EM	FDB	Happy, sad, neutral	Adaptive Functioning	VABS	0.37 (0.120)
Wallace et al. (2011)	42 (38)	ASD/AS/ PDD	15.7	EL	EFAS (morphed)	Basic Emotions	Adaptive Functioning	ABAS	0.42 (0.005)
Williams and Gray (2013a)	42 (38)	ASD	5.3	EL/ EM	EFAS	Basic Emotions	Adaptive Functioning	VABS	0.21 (0.192)
Williams and Gray (2013b)	55 (48)	ASD	5.2	EL/ EM	EFAS	Basic Emotions	Adaptive Functioning	VABS	0.36 (0.007)
Buitelaar and van der Wees (1997)	20 (18)	PDD	12.5	EL	EFAS	Basic Emotions	ToM	10 first and second-order ToM tasks.	0.40 (0.081)
Buitelaar, van der Wees, Swaab-Barneveld, and Gaag, (1999) – ASD	20 (18)	ASD	12.5	EM	EFAS	Basic Emotions	ToM	10 first and second-order ToM tasks.	0.43 (0.058)
Buitelaar et al. (1999) – PDD	20 (17)	PDD	12.4	EM	EFAS	Basic Emotions	ToM	10 first and second-rder ToM tasks.	0.37 (0.109)
Dyck et al. (2006) – EXP. #2	30 (23)	ASD	8.5	EL	FCT	Basic Emotions	ToM	3 first and second-order tasks	0.74 (<0.001)
Heerey et al. (2003)	25 (20)	ASD/AS	10.7	EL	NR	Basic emotions + Embarrassment/ shame	ToM	Strange Stories Task	0.49 (0.012)
Rice et al. (2015)	31 (28)	ASD	7.7	EL	NEPSY-II (affect rec. subtest)	NR	ToM	NEPSY-II (ToM Subtest)	0.52 (0.002)
Cook, Brewer, Shah, and Bird (2013) – EXP. #1	16 (15)	ASD	39.2	Other	EFAS (morphed)	Surprise-fear/disgust-anger continua	Alexithymia	TAS-20	−0.46 (0.075)
Kätsyri et al. (2008)	20 (13)	AS	19.5	EL	EFAS/TKK Collection	Anger, disgust, fear, happy	Alexithymia	TAS-20	−0.21 (0.634)
	29 (0)	ASD	41.35	EL	EFAS	Basic Emotions	Alexithymia	BVAQ	

(*Continued*)

Study	N (males)	Diagnosis	Mean age (years)	FER Type	FER Stimuli	Emotions measured	Variable Category	Measure	Pearson's *r* estimate (p)
Ketelaars et al. (2016)									−0.36 (0.055)
Kliemann, Rosenblau, Bölte, Heekeren, and Dziobek (2013)	24 (15)	ASD/AS	30.4	EL/EM	Face Puzzle	NR	Alexithymia	TAS-26 (German)	−0.27 (0.213)
Braverman et al. (1989)	22 (NR)	ASD/PDD	10.8	EM	EFAS	Basic Emotions	Verbal Intelligence	SBIS/EPVT	0.48 (0.022)
Buitelaar et al. (1999) – ASD	20 (18)	ASD	12.5	EM	EFAS	Basic Emotions	Verbal Intelligence	WISC	0.34 (0.144)
Buitelaar et al. (1999) – PDD	20 (17)	PDD	12.4	EM	EFAS	Basic Emotions	Verbal Intelligence	WISC	0.27 (0.254)
Dyck et al. (2006) – EXP. #2	30 (23)	ASD	8.5	EL	FCT	Basic Emotions	Verbal Intelligence	WISC	0.57 (0.001)
Humphreys et al. (2007)	20 (19)	ASD	24.0	EL	FEEST (morphed)	Basic Emotions	Verbal Intelligence	WISC	0.52 (0.017)
Kätsyri et al. (2008)	20 (13)	AS	19.5	EL	EFAS/TKK Collection	Anger, disgust, fear, happy	Verbal Intelligence	WAIS	0.36 (0.399)
Loveland, Bachevalier, Pearson, and Lane (2008)	80 (NR)	ASD	12	EM	NR	NR	Verbal Intelligence	SBIS	0.52 (<0.001)
Ozonoff et al. (1991)	23 (21)	ASD/PDD	12.1	EM	IST	Basic and Complex	Verbal Intelligence	WISC/WAIS	0.28 (0.198)
Teunisse and de Gelder (2001)	17 (13)	ASD	21.3	EL	EFAS (morphed)	Anger-sad/anger-fear/happy-sad continua	Verbal Intelligence	'woordenlijst' (wordlist) of GIT	0.39 (0.123)
Wallace, Coleman, and Bailey (2008) – EXP. #1	26 (24)	ASD/AS	32.0	EL	EFAS (upright and inverted)	Basic Emotions	Verbal Intelligence	BPVS	0.47 (0.014)
Williams and Gray (2013a)	42 (38)	ASD	5.3	EL/EM	EFAS	Basic Emotions	Verbal Intelligence	WPPSI	0.15 (0.349)
Williams and Gray (2013b)	55 (48)	ASD	5.2	EL/EM	EFAS	Basic Emotions	Verbal Intelligence	WPPSI	0.40 (0.002)
Braverman et al. (1989)	22 (NR)	ASD/PDD	10.8	EM	EFAS	Basic Emotions	Nonverbal Intelligence	PPVT	0.57 (0.005)
Buitelaar et al. (1999) – ASD	20 (18)	ASD	12.5	EM	EFAS	Basic Emotions	Nonverbal Intelligence	WISC	0.39 (0.090)
Buitelaar et al. (1999) – PDD	20 (17)	PDD	12.4	EM	EFAS	Basic Emotions	Nonverbal Intelligence	WISC	0.24 (0.313)
Dyck et al. (2006) – EXP. #2	30 (23)	ASD	8.5	EL	FCT	Basic Emotions	Nonverbal Intelligence	WISC	0.26 (0.167)
Humphreys et al. (2007)	20 (19)	ASD	24.0	EL	FEEST (morphed)	Basic Emotions	Nonverbal Intelligence	WISC	0.49 (0.027)
Kätsyri et al. (2008)	20 (13)	AS	19.5	EL	EFAS/TKK Collection	Anger, disgust, fear, happy	Nonverbal Intelligence	WAIS	0.33 (0.443)
Ozonoff et al. (1991)	23 (21)	ASD/PDD	12.1	EM	IST	Basic and Complex	Nonverbal Intelligence	WAIS/WISC	0.21 (0.340)
Tantam et al. (2006)	10 (8)	ASD	12.1	EL/Other	EFAS	Basic Emotions	Nonverbal Intelligence	Raven's Matrices	0.57 (0.087)
Teunisse and de Gelder (2001)	17 (13)	ASD	21.3	EL	EFAS (morphed)	Anger-sad/anger-fear/happy-sad continua	Nonverbal Intelligence	Raven's Matrices	0.32 (0.213)
Wallace et al. (2008) – EXP. #1	26 (24)	ASD/AS	32.0	EL	EFAS (upright and inverted)	Basic Emotions	Nonverbal Intelligence	Raven's Matrices	0.22 (0.283)
Williams and Gray (2013b)	55 (48)	ASD	5.2	EL/EM	EFAS	Basic Emotions	Nonverbal Intelligence	WPPSI	0.38 (0.004)
Bal et al. (2010)	14 (NR)	ASD/PDD	10.3	EL	DARE (morphed)	Basic Emotions	ASD Symptomology	SRS	−0.54 (0.045)
Cook et al. (2013) – EXP. #1	16 (15)	ASD	39.2	Other	EFAS (morphed)	Surprise-fear/disgust-anger continua	ASD Symptomology	AQ	−0.20 (0.460)
Evers, Steyaert, Noens, & Wagemans (2015)	50 (47)	ASD	10.8	EL	ERT	Basic Emotions	ASD Symptomology	SRS	−0.17 (0.239)
Grove, Baillie, Allison, Baron-Cohen, & Hoekstra (2014)	363 (193)	ASD	36.0	EL	RTEM/KDEF	Complex/Basic Emotions	ASD Symptomology	AQ	−0.32 (<0.001)
	20 (19)	ASD	24.0	EL		Basic Emotions		ADOS/SRS	

(*Continued*)

Study	N (males)	Diagnosis	Mean age (years)	FER Type	FER Stimuli	Emotions measured	Variable Category	Measure	Pearson's *r* estimate (p)
Humphreys et al. (2007)					FEEST (morphed)		ASD Symptomology		−0.52 (0.017)
Kliemann et al. (2013)	24 (15)	ASD/AS	30.4	EL/ EM	Face Puzzle	NR	ASD Symptomology	ADOS/SRS	−0.55 (0.005)
Ozonoff et al. (1991)	23 (21)	ASD/PDD	12.1	EM	IST	Basic and Complex	ASD Symptomology	CARS	−0.25 (0.253)
Rice et al. (2015)	31 (28)	ASD	7.7	EL	NEPSY-II (affect rec. subtest)	NR	ASD Symptomology	SRS	−0.53 (0.002)
Tanaka et al. (2012)	55 (NR)	ASD/AS/ PDD	12.0	EM	Matchmaker Expression Stimuli	Basic emotions	ASD Symptomology	ADOS	−0.39 (0.001)
Tardif et al. (2007)	12 (9)	ASD	10.4	EM	Research-created stimuli/task	Happy, sad, surprise, disgust	ASD Symptomology	CARS	−0.77 (0.002)
Wallace et al. (2011)	42 (38)	ASD/AS/ PDD	15.7	EL	EFAS (morphed)	Basic Emotions	ASD Symptomology	ADOS	−0.16 (0.323)
Williams & Gray (2013a)	42 (38)	ASD	5.3	EL/ EM	EFAS	Basic Emotions	ASD Symptomology	ADOS	−0.17 (0.295)
Williams & Gray (2013b)	55 (48)	ASD	5.2	EL/ EM	EFAS	Basic Emotions	ASD Symptomology	ADOS/SRS	−0.33 (0.013)
Buitelaar et al. (1999) – ASD	20 (18)	ASD	12.5	EM	EFAS	Basic Emotions	Age	CA	0.30 (0.202)
Buitelaar et al. (1999) – PDD	20 (17)	PDD	12.4	EM	EFAS	Basic Emotions	Age	CA	0.41 (0.072)
Gepner et al. (2001)	13 (12)	ASD	5.8	EM	Researcher-created stimuli	Happy, surprise, sad, disgust	Age	CA	0.19 (0.543)
Gepner, Deruelle, and Grynfeltt (1996)	7 (4)	ASD	11.3	EL	Researcher-created stimuli	Happy, surprise, dislike?, neutral	Age	CA	−0.03 (0.952)
Kätsyri et al. (2008)	20 (13)	AS	19.5	EL	EFAS/TKK Collection	Anger, disgust, fear, happy	Age	CA	0.24 (0.584)
Law Smith, Montagne, Perrett, Gill, and Gallagher (2010)	21 (21)	ASD/AS	15.3	EL	ERT (varying thresholds)	Basic Emotions	Age	CA	0.53 (0.012)
Ozonoff et al. (1991)	23 (21)	ASD/PDD	12.1	EM	IST	Basic and Complex	Age	CA	0.04 (0.858)
Rump, Giovannelli, Minshew, and Strauss (1980)	19 (14)	ASD	6.4	EL	Researcher-created stimuli	Happy, sad, angry, fear	Age	CA	0.33 (0.170)
Tanaka et al. (2012)	85 (NR)	ASD/AS/ PDD	12.0	EM	Matchmaker Expression Stimuli	Basic emotions	Age	CA	0.29 (0.007)
Williams and Gray (2013a)	42 (38)	ASD	5.3	EL/ EM	EFAS	Basic Emotions	Age	CA	0.20 (0.210)
Williams and Gray (2013b)	55 (48)	ASD	5.2	EL/ EM	EFAS	Basic Emotions	Age	CA	0.32 (0.017)

Note. NR = Not Reported. Some correlation coefficients in this table were not reported directly in the original papers, but estimated from *p*-values or other statistics using meta-analytic conversion formulas. Some coefficients in the table were averaged using Fisher's Z transformations if authors reported correlations between FER and two or more measures relevant to the same variable category. The *p*-values reported in this table are estimated by CMA, and not consistent with *p*-values in original reports.

Acronyms. ABAS = Adaptive Behavior Assessment System, ADOS = Autism Diagnostic Observation Scale, AQ = Autism Spectrum Quotient, Basic Emotions = Happiness, anger, sadness, fear, disgust, surprise, AS = Asperger's Syndrome; ASD = Autistic Disorder; BPVS = British Picture Vocabulary Scale, BVAQ = Bermond Vorst Alexithymia Questionnaire, CARS = Childhood Autism Rating Scale, CATS = Comprehensive Affective Testing System, Complex Emotions = See, Baron-Cohen, Wheelwright, Hill, Raste & Plumb (2001) for a full list, DANVA = Diagnostic Analysis of Nonverbal Accuracy, DARE = Dynamic Affect Recognition Evaluation, EFAS = Ekman Facial Affect Set, EPVT = Expressive One-Word Picture Vocabulary Test, ERT = Emotion Recognition Task, FCT = Facial Cues Test, FDB = Facial Discrimination Battery, FEEST = Facial Expressions of Emotion: Stimuli and Tests, GIT = Groninger Itelligentie Test, IST = Izard Stimuli Set, KDEF = Karolinska Directed Emotional Faces, NEPSY = A Developmental NEuroPSYchological Assessment, PDD = Pervasive Developmental

Disorder; RTEM = Reading the Mind in the Eyes Test, SBIS = Stanford-Binet Intelligence Scales, SRS = Social Responsiveness Scale, TAS = Toronto Alexithymia Questionnaire, VABS = Vineland Adaptive Behavior Scales, WISC = Wechsler Intelligence Scale for Children, WPPSI = Wechsler Preschool and Primary Scale of Intelligence.

References[1]

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

Bachevalier, J., & Loveland, K. A. (2006). The orbitofrontal-amygdala circuit and self-regulation of social-emotional behaviour in autism. *Neuroscience and Biobehavioral Reviews*, *30*, 97–117. http://dx.doi.org/10.1016/j.neubiorev.2005.07.002.

Bagby, R. M., Parker, J. D., & Taylor, G. J. (1994). The twenty-item Toronto Alexithymia Scale—I. Item selection and cross-validation of the factor structure. *Journal of Psychosomatic Research*, *38*(1), 23–32.

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

Baron-Cohen, S., & Wheelwright, S. (2003). The Friendship Questionnaire: An investigation of adults with Asperger syndrome or high-functioning autism, and normal sex differences. *Journal of Autism and Developmental Disorders*, *33*(5), 509–517.

Baron-Cohen, S., Leslie, A., & Frith, U. (1985). Does the autistic child have a 'theory of mind'? *Cognition*, *21*, 37–46.

Baron-Cohen, S., Wheelwright, S., Hill, J., Raste, Y., & Plumb, I. (2001). The Reading the Mind in the Eyes test revised version: A study with normal adults, and adults with Asperger syndrome or high-functioning autism. *Journal of Child Psychology, and Psychiatry*, *42*(2), 241–251.

Baron-Cohen, S. ([257_TD$DIFF]2005). The empathizing system: A revision of the 1994 model of the mind-reading system. In B. Ellis, & D. Bjorklund (Eds.), *Origins of the social mind* (pp. 468–490).New York: Guildford Press.

Baron-Cohen, S. (1994). The mindreading system: New directions for research. *Current Psychology of Cognition*, *13*(5), 724–750.

Bermond, B. (1998). *Bermond-vorst alexithymia questionnaire. Unpublished document*. The Netherlands: University of Amsterdam.

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.

Birmingham, E., Ristic, J., & Kingstone, A. (2012). Investigating social attention: A case for increasing stimulus complexity in the laboratory. In J. A. Burack, J. T. Enns, & N. A. Fox (Eds.), *Cognitive science, development, and psychopathology: typical and atypical developmental trajectories of attention* (pp. 251–276).New York, NY: Oxford University Press.

Birmingham, E., Meixner, T., Iarocci, G., Kanan, C., Smilek, D., & Tanaka, J. (2013). The Moving Window Technique: A window into developmental changes in attention during facial emotion recognition. *Child Development*, *84*(4), 1407–1424. http://dx.doi.org/10.1111/cdev.12039.Epub 2012Dec17.

Borenstein, M., Hedges, L., Higgins, J., & Rothstein, H. (2005). *Comprehensive meta-analysis version 2*. Englewood, NJ: Biostat.

Borenstein, M., Hedges, L., & Rothstein, D. (2007). *Meta-analysis: fixed effect vs random effects*. Retrieved from http://www.meta-analysis.com/downloads/Meta-analysis_fixed_effect_vs_ random_effects_sv.pdf.

Borenstein, M., Hedges, L. V., Higgins, J. P. T., & Rothstein, H. R. (2009). *Introduction to meta-analysis*. Chichester, UK: Wiley.

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

Buitelaar*, J. K., & van der Wees, M. (1997). Are deficits in the decoding of affective cues and in mentalizing abilities independent? *Journal of Autism and Developmental Disorders*, *27*(5), 539–556.

Buitelaar*, J. K., van der Wees, M. V. D., Swaab-Barneveld, H., & Gaag, R. J. V. D. (1999). Verbal memory and performance IQ predict theory of mind and emotion recognition ability in children with autistic spectrum disorders and in psychiatric control children. *Journal of Child Psychology and Psychiatry*, *40*(6), 869–881.

Čeponienė, R., Lepistö, T., Shestakova, A., Vanhala, R., Alku, P., Näätänen, R., et al. (2003). Speech-sound-selective auditory impairment in children with autism: They can perceive but do not attend. *Proceedings of the National Academy of Sciences of the United States of America*, *100*(9), 5567–5572.

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.

Constantino, J. N., & Gruber, C. P. (2002). *The social responsiveness scale*. Los Angeles: Western Psychological Services.

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 EXP. #1.

Cronbach, L. J., & Meehl, P. E. (1955). Construct validity in psychological tests. *Psychological Bulletin*, *51*, 282–302.

Cumming, G. (2013). The new statistics why and how. *Psychological Science*, *25*(1), 7–29.

DeThorne, L. S., & Schaefer, B. A. (2004). A guide to child nonverbal IQ measures. *American Journal of Speech-Language Pathology*, *13*(4), 275–290.

Duval, S., & Tweedie, R. (2000). Trim and fill: A simple funnel-plot-based method of testing and adjusting for publication bias in meta-analysis. *Biometrics*, *56*(2), 455–463.

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 EXP. #2.

Ekman, P., Friesen, W. V., & Hager, J. C. (2002). *Facial action coding system: the manual on CD-ROM*. Salt Lake City, UT: Instructor's Guide.

Evers*, K., Steyaert, J., Noens, I., & Wagemans, J. (2015). Reduced recognition of dynamic facial emotional expressions and emotion-specific response bias in children with an autism spectrum disorder. *Journal of Autism and Developmental Disorders*, *45*(6), 1774–1784.

Frith, U. (2003). *Autism: explaining the enigma*, 2nd ed. Oxford: Blackwell Publishing.

Furr, R. M., & Bacharach, V. R. (2013). *Psychometrics: an introduction*. Thousand Oaks, California: Sage.

Gaigg, S. B. (2012). The interplay between emotion and cognition in autism spectrum disorder: Implications for developmental theory. *Frontiers in Integrative Neuroscience*, *6*, 1–35.

García-Villamisar*, D., Rojahn, J., Zaja, R. H., & Jodra, M. (2010). Facial emotion processing and social adaptation in adults with and without autism spectrum disorder. *Research in Autism Spectrum Disorders*, *4*(4), 755–762.

Gepner*, B., Gelder, B. D., & Schonen, S. D. (1996). Face processing in autistics: Evidence for a generalised deficit? *Child Neuropsychology*, *2*(2), 123–139.

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.

Golan, O., & Baron-Cohen, S. (2006). Systemizing empathy: Teaching adults with Asperger syndrome or high-functioning autism to recognize complex emotions using interactive multimedia. *Development and Psychopathology*, *18*(02), 591–617.

Golan, O., Baron-Cohen, S., Hill, J. J., & Golan, Y. (2006). The reading the mind in films task: Complex emotion recognition in adults with and without autism spectrum conditions. *Social Neuroscience*, *1*(2), 111–123.

Golan, O., Baron-Cohen, S., & Golan, Y. (2008). The 'reading the mind in films' task [child version]: Complex emotion and mental state recognition in children with and without autism spectrum conditions. *Journal of Autism and Developmental Disorders*, *38*(8), 1534–1541.

Golan, O., Ashwin, E., Granader, Y., McClintock, S., Day, K., Leggett, V., et al. (2010). Enhancing emotion recognition in children with autism spectrum conditions: An intervention using animated vehicles with real emotional faces. *Journal of Autism and Developmental Disorders*, *40*(3), 269–279.

Goldman, A. I., & Sripada, C. S. (2005). Simulationist models of face-based emotion recognition. *Cognition*, *94*(3), 193–213.

[1] Note. Articles used for meta-analysis are marked with an asterisk (*).

Grove, R., Baillie, A., Allison, C., Baron-Cohen, S., & Hoekstra, R. A. (2014). The latent structure of cognitive and emotional empathy in individuals with autism, first-degree relatives and typical individuals. *Molecular Autism*, *5*(1), 42.

Grynberg, D., Chang, B., Corneille, O., Maurage, P., Vermeulen, N., Berthoz, S., et al. (2012). Alexithymia and the processing of emotional facial expressions (EFEs): systematic review, unanswered questions and further perspectives. *PLoS One*, *7*(8), e42429.

Hall, J. A., Andrzejewski, S. A., & Yopchick, J. E. (2009). Psychosocial correlates of interpersonal sensitivity: A meta-analysis. *Journal of Nonverbal Behavior*, *33* (3), 149–180.

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.

Harrison, P., & Oakland, T. (2003). *Adaptive behavior assessment system (ABAS-II)*. San Antonio, TX: The Psychological Corporation.

Heerey*, E. A., Keltner, D., & Capps, L. M. (2003). Making sense of self-conscious emotion: Linking theory of mind and emotion in children with autism. *Emotion*, *3*(4), 394.

Henson, R. K. (2006). Effect-size measures and meta-analytic thinking in counseling psychology research. *The Counseling Psychologist*, *34*(5), 601–629.

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

Hunter, J. E., & Schmidt, F. L. (2004). *Methods of meta-analysis: correcting error and bias in research findings*. New York: Sage Publications.

Izard, C., Fine, S., Schultz, D., Mostow, A., Ackerman, B., & Youngstrom, E. (2001). Emotion knowledge as a predictor of social behavior and academic competence in children at risk. *Psychological Science*, *12*(1), 18–23.

Jones, C. R., Pickles, A., Falcaro, M., Marsden, A. J., Happé, F., Scott, S. K., . . . Simonoff, E. (2011). A multimodal approach to emotion recognition ability in autism spectrum disorders. *Journal of Child Psychology and Psychiatry*, *52*(3), 275–285.

Kätsyri*, J., Saalasti, S., Tiippana, K., von Wendt, L., & Sams, M. (2008). Impaired recognition of facial emotions from low-spatial frequencies in Asperger syndrome. *Neuropsychologia*, *46*(7), 1888–1897.

Kandalaft, M. R., Didehbani, N., Krawczyk, D. C., Allen, T. T., & Chapman, S. B. (2013). Virtual reality social cognition training for young adults with high-functioning autism. *Journal of Autism and Developmental Disorders*, *43*(1), 34–44.

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.

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*(76), 1–13.

Klin, A., Jones, W., Schultz, R., Volkmar, F., & Cohen, D. (2002). Visual fixation patterns during viewing of naturalistic social situations as predictors of social competence in individuals with autism. *Archives of General Psychiatry*, *59*(9), 809–816.

Klin, A. (1991). Young autistic children's listening preferences in regard to speech: A possible characterization of the symptom of social withdrawal. *Journal of Autism and Developmental Disorders*, *21*(1), 29–42.

Kupfer, J., Brosig, B., & Brähler, E. (2001). *Toronto-Alexithymie-Skala-26 (TAS-26)*. Göttingen: Hogrefe.

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*, 2777–2781.

Lord, C., Rutter, M., & Le Couteur, A. (1994). Autism diagnostic interview-revised: A revised version of a diagnostic interview for caregivers of individuals with possible pervasive developmental disorders. *Journal of Autism and Developmental Disorders*, *24*(5), 659–685.

Lord, C., Risi, S., Lambrecht, L., Cook Jr, E. H., Leventhal, B. L., DiLavore, P. C., et al. (2000). The autism diagnostic observation schedule-generic: A standard measure of social and communication deficits associated with the spectrum of autism. *Journal of Autism and Developmental Disorders*, *30*(3), 205–223.

Loveland, K. A., & Tunali, B. (1991). Social scripts for conversational interactions in autism and Down syndrome. *Journal of Autism and Developmental Disorders*, *21*(2), 177–186.

Loveland*, K. A., Bachevalier, J., Pearson, D. A., & Lane, D. M. (2008). Fronto-limbic functioning in children and adolescents with and without autism. *Neuropsychologia*, *46*(1), 49–62.

Loveland, K. A. (2001). Toward an ecological theory of autism. In J. A. Burack, T. Charman, N. Yirmiya, & P. R. Zelazo (Eds.), *The development of autism: perspectives from theory and research* (pp. 17–37).Mahwah, NJ: Lawrence Erlbaum Associates.

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 pt. 1), 933–945.

Nemiah, J. C., Freyberger, H., & Sifneos, P. E. (1976). Alexithymia: A view of the psychosomatic process. *Modern Trends in Psychosomatic Medicine*, *3*, 430–439.

Nowicki, S., & Duke, M. (2013). Accuracy in interpreting nonverbal cues. In J. A. Hall, & M. L. Knapp (Eds.), *Nonverbal communication* (pp. 441–470).Boston: De Gruyter Mouton.

Ozonoff*, S., Rogers, S. J., & Pennington, B. F. (1991). Asperger's syndrome: Evidence of an empirical distinction from high-functioning autism. *Journal of Child Psychology and Psychiatry*, *32*(7), 1107–1122.

Perner, J., Frith, U., Leslie, A. M., & Leekam, S. R. (1989). Exploration of the autistic child's theory of mind: Knowledge, belief, and communication. *Child Development*689–700.

Premack, D. G., & Woodruff, G. (1978). Does the chimpanzee have a theory of mind? *Behavioral and Brain Sciences*, *1*(4), 515–526.

Prior, M., Dahlstrom, B., & Squires, T. L. (1990). Autistic children's knowledge of thinking and feeling states in other people. *Journal of Child Psychology and Psychiatry*, *31*(4), 587–601.

Reichow, B., & Volkmar, F. R. (2010). Social skills interventions for individuals with autism: Evaluation for evidence-based practices within a best evidence synthesis framework. *Journal of Autism and Developmental Disorders*, *40*(2), 149–166.

Riby, D. M., & Hancock, P. J. B. (2008). Viewing it differently: Social scene perception in Williams syndrome and autism. *Neuropsychologia*, *46*, 2855–2860.

Rice*, L. M., Wall, C. A., Fogel, A., & Shic, F. (2015). Computer-assisted face processing instruction improves emotion recognition, mentalizing, and social skills in students with ASD. *Journal of Autism and Developmental Disorders*1–11.

Rieffe, C., Terwogt, M. M., & Stockmann, L. (2000). Understanding atypical emotions among children with autism. *Journal of Autism and Developmental Disorders*, *30*(3), 195–203.

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

Roid, G., & Barram, R. (2004). *Essentials of Stanford–Binet intelligence scales (SB5) assessment*. Hoboken, New Jersey: John Wiley & Sons, Inc..

Rosenthal, R. (1991). *Meta-analytic procedures for social research*, revised ed. Newbury Park, CA: Sage.

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.

Schopler, E., Reichler, R. J., DeVellis, R. F., & Daly, K. (1980). Toward objective classification of childhood autism: Childhood autism rating scale (CARS). *Journal of Autism and Developmental Disorders*, *10*(1), 91–103.

Sigman, M. D., Kasari, C., Kwon, J. H., & Yirmiya, N. (1992). Responses to the negative emotions of others by autistic, mentally retarded, and normal children. *Child Development*796–807.

Silver, M., & Oakes, P. (2001). Evaluation of a new computer intervention to teach people with autism or Asperger syndrome to recognize and predict emotions in others. *Autism*, *5*(3), 299–316.

Sparrow, S. S., Balla, D. A., & Cicchetti, D. V. (1984). *Vineland adaptive behavior scales*. Circle Pine, MN: American Guidance Services.

Sparrow, S. S., Cicchetti, D. A., & Balla, D. V. (2005). *Vineland adaptive behavior scales*, 2nd ed. San Antonio, TX: The Psychological Corporation.

Tanaka*, J. W., Wolf, J. M., Klaiman, C., Koenig, K., Cockburn, J., Herlihy, L., et al. (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*, *53*(12), 1259–1267.

Tantam*, D., Monaghan, L., Nicholson, H., & Stirling, J. (2006). Autistic children's ability to interpret faces: A research note. *Journal of Child Psychology and Psychiatry*, *30*(4), 623–630.

Tardif*, C., Lainé, F., Rodriguez, M., & Gepner, B. (2007). Slowing down presentation of facial movements and vocal sounds enhances facial expression recognition and induces facial–vocal imitation in children with autism. *Journal of Autism and Developmental Disorders*, *37*(8), 1469–1484.

Taylor, G. J., Bagby, R. M., & Parker, J. D. A. (1992). The revised Toronto alexithymia scale: Some reliability, validity, and normative data. *Psychotherapy and Psychosomatics*, *57*(1–2), 34–41.

Taylor, G. J., Bagby, R. M., & Parker, J. D. (2003). The 20-item Toronto alexithymia scale: IV. Reliability and factorial validity in different languages and cultures. *Journal of Psychosomatic Research*, *55*(3), 277–283.

Teunisse*, J. P., & de Gelder, B. (2001). Impaired categorical perception of facial expressions in high-functioning adolescents with autism. *Child Neuropsychology*, *7*(1), 1–14.

Uljarevic, M., & Hamilton, A. (2013). Recognition of emotions in autism: A formal meta-analysis. *Journal of Autism and Developmental Disorders*, *43*(7), 1517–1526.

Valentine, J. C., Pigott, T. D., & Rothstein, H. R. (2010). How many studies do you need? A primer on statistical power for meta-analysis. *Journal of Educational and Behavioral Statistics*, *35*(2), 215–247.

Volkmar, F. R., Sparrow, S. S., Goudreau, D., Cicchetti, D. V., Paul, R., & Cohen, D. J. (1987). Social deficits in autism: An operational approach using the Vineland Adaptive Behavior Scales. *Journal of the American Academy of Child & Adolescent Psychiatry*, *26*(2), 156–161.

Wallace*, S., Coleman, M., & Bailey, A. (2008). An investigation of basic facial expression recognition in autism spectrum disorders. *Cognition & Emotion*, *22*(7), 1353–1380 EXP. #1.

Wallace*, G. L., Case, L. K., Harms, M. B., Silvers, J. A., Kenworthy, L., & Martin, A. (2011). Diminished sensitivity to sad facial expressions in high functioning autism spectrum disorders is associated with symptomatology and adaptive functioning. *Journal of Autism and Developmental Disorders*, *41*(11), 1475–1486.

Wechsler, D. (1991). *WISC-III: Wechsler intelligence scale for children: manual*. San Antonio, TX: Psychological Corporation.

Wechsler, D. (2008). *Wechsler adult intelligence scale*, Fourth ed. (WAIS–IV) San Antonio, TX: NCS Pearson.

Williams*, B. T., & Gray, K. M. (2013a). Are emotion recognition skills related to autism symptom severity in children with autism? *International Journal of Developmental Disabilities*, *59*(2), 118–133.

Williams*, B. T., & Gray, K. M. (2013b). The relationship between emotion recognition ability and social skills in young children with autism. *Autism*, *17*(6), 762–768.

Williams, D. (2010). Theory of own mind in autism: Evidence of a specific deficit in self-awareness? *Autism*, *14*(5), 474–494.

Woodbury-Smith, M. R., Robinson, J., Wheelwright, S., & Baron-Cohen, S. (2005). Screening adults for Asperger syndrome using the AQ: A preliminary study of its diagnostic validity in clinical practice. *Journal of Autism and Developmental Disorders*, *35*(3), 331–335.

Yager, J., & Iarocci, G. (2013). The development of the multidimensional social competence scale: A standardized measure of social competence in autism spectrum disorders. *Autism Research*, *6*(6), 631–641.