Clinical Psychology Review 120 (2025) 102623

Contents lists available at ScienceDirect

Clinical Psychology Review

journal homepage: www.elsevier.com/locate/clinpsychrev

Review

A systematic review and meta-analysis of empathy in autism: The influence of measures

Noémie M. Cusson [a,b], Alexa J. Meilleur [a,b], Boris C. Bernhardt [c], Isabelle Soulières [a,b], Laurent Mottron [b,d,*]

[a] Department of Psychology, Université du Québec à Montréal, Pavillon Adrien-Pinard, 100 Rue Sherbrooke O., Montréal, Québec H2X 3P2, Canada
[b] CIUSSS-NIM Research Centre, Rivière-des-Prairies Hospital, 7070 Boulevard Perras, Montréal, Québec H1E 1A4, Canada
[c] McConnell Brain Imaging Centre, Montréal Neurological Institute and Hospital, McGill University, 3801 University Street, Montréal, Québec H3A 2B4, Canada
[d] Department of Psychiatry and Addictology, Université de Montréal, Pavillon Roger-Gaudry, 2900 Boulevard Édouard-Montpetit, Montréal, Québec H3T 1J4, Canada

A R T I C L E I N F O

Keywords:
Autism
Empathy
Empathy measurement
Meta-analysis
Systematic review

A B S T R A C T

Empathy deficits are considered a core attribute of autism and are scored in standardized autism diagnostic instruments. However, empirical evidence concerning empathy in autism is contradictory. This systematic review, which included 226 studies, thus offers a comprehensive overview of empathy in autism. It additionally examined the impact of the chosen empathy measure and the effect of several moderators. The results reveal a large effect size for cognitive empathy ($g = -0.85$) and unidimensional empathy ($g = -1.70$), but only a small effect size for affective empathy ($g = -0.17$), which became non-significant when limiting analyses to high-quality studies. Meta-regressions suggest that publication year, study quality, alexithymia, verbal IQ, and age do not moderate empathy, whereas sex specifically moderates unidimensional empathy. Critically, there were notable differences in effect sizes obtained across empathy measures and even between subscales of the same measure. For instance, results for the affective empathy subscales of the Interpersonal Reactivity Index reveal lower empathic concern ($g = -0.59$) but increased personal distress ($g = 0.67$) in autistic relative to typical participants. A qualitative review of ecological and neuroimaging tasks mostly demonstrated minimal autistic versus non-autistic differences. This meta-analysis thus suggests that measuring empathy as a unidimensional construct may both distort and increase the notion of an empathy deficit in autism.

1. Introduction

Empathy plays an essential role in typical social interactions, as well as in altruistic and moral behaviour (e.g., de Waal & Preston, 2017; McDonald & Messinger, 2011). Autism has been conceptualized as a disorder of empathy due to the purported difficulties of autistic individuals with theory of mind (Gillberg, 1992) or with socio-emotional reciprocity and socio-communicative behaviours (Henry, von Hippel, Molenberghs, Lee, & Sachdev, 2016; Preston & de Waal, 2002). However, the magnitude and nature of these presumed empathy difficulties warrant a thorough review before autism can be considered as an “empathy disorder”.

1.1. Empathy conceptualization and definition

Empathy is a multidimensional construct, with both cognitive and affective aspects. However, several conceptualizations of empathy coexist, spanning across psychology, cognitive neuroscience, philosophy, anthropology and literary studies (Bernhardt & Singer, 2012; Engelen & Röttger-Rössler, 2012). A literature review identified 31 separate empathy definitions in autism research, which varied according to 12 dimensions, and 52 different methods to assess empathy (Bollen, 2023), whereas another wider review found 43 unique empathy conceptualizations, which diverged according to eight dimensions (Cuff, Brown, Taylor, & Howat, 2016). The latter review further suggested a new empathy definition in the light of the identified dimensions:

> Empathy is an emotional response (affective), dependent upon the interaction between trait capacities and state influences. Empathic processes are automatically elicited but are also shaped by top-down control processes. The resulting emotion is similar to one’s perception (directly experienced or imagined) and understanding

* Corresponding author at: Rivière-des-Prairies Hospital, 7070 Boulevard Perras, Montréal, Québec H1E 1A4, Canada.
E-mail address: laurent.mottron@umontreal.ca (L. Mottron).

https://doi.org/10.1016/j.cpr.2025.102623
Received 16 July 2024; Received in revised form 12 May 2025; Accepted 22 July 2025
Available online 24 July 2025
0272-7358/

(cognitive empathy) of the stimulus emotion, with recognition that the source of the emotion is not one's own. (p. 150, Cuff et al., 2016).

This definition thus highlights that empathy has both cognitive and affective (or emotional) components, which is consistent with findings from the field of neuroscience. Meta-analyses assessing the neural basis of empathy indeed suggest the existence of a network that is consistently recruited in empathy tasks, in addition to specific activations in distinct regions underlying either cognitive or affective empathy (Fan, Duncan, de Greck, & Northoff, 2011; Kogler, Muller, Werminghausen, Eickhoff, & Derntl, 2020; Lamm, Decety, & Singer, 2011).

This definition also allows to further distinguish empathy from related concepts. For instance, it highlights that empathy requires some degree of self-other distinction, that is, the ability to discern whether the emotion that one feels is caused by another person or arises from within oneself (Cuff et al., 2016; Singer & Lamm, 2009). This is what, according to several authors, differentiates empathy from emotional contagion (i.e., the tendency to automatically have a matching emotional state in response to perceiving another person's emotional state), which does not entail self-other distinction (Bernhardt & Singer, 2012; Bird & Viding, 2014). However, other authors nevertheless consider emotional contagion as a facet or as a basic form of empathy (e.g., de Waal & Preston, 2017; Zaki & Ochsner, 2012).

Similarly, cognitive empathy has sometimes been used synonymously with theory of mind (e.g., Lawrence, Shaw, Baker, Baron-Cohen, & David, 2004; McDonald & Messinger, 2011). Indeed, *theory of mind* can be defined as the ability to understand another person's thoughts, intentions, beliefs, and feelings (Dvash & Shamay-Tsoory, 2014). It can also be separated into cognitive and affective components. The cognitive component refers to the ability to understand another person's thoughts, intentions, and beliefs, that is, their cognitive state, and the affective component, to the ability to understand another person's affective state (Henry et al., 2016; Walter, 2012). Therefore, only the affective component of theory of mind appears to be synonymous with cognitive empathy.

1.2. Measuring empathy

Given the existence of multiple and sometimes conflicting empathy definitions, empathy has been measured in a number of ways. When measured as a unidimensional construct, a single empathy score is obtained. When measured as a multidimensional construct, multiple empathy components are assessed and separate scores are calculated for each component. Empathy is frequently measured with a self-report questionnaire, two of the most common being the Empathy Quotient (EQ; Baron-Cohen & Wheelwright, 2004) and the Interpersonal Reactivity Index (IRI; Davis, 1980). The EQ comprises 40 items that assess empathy and 20 filler items (Baron-Cohen & Wheelwright, 2004). As it allows the calculation of a single total empathy score, it is considered to be a unidimensional empathy measure (but see Allison, Baron-Cohen, Wheelwright, Stone, & Muncer, 2011; Lawrence et al., 2004). The IRI is a questionnaire containing 28 items separated into four seven-item subscales, namely, perspective taking (PT), fantasy (F), empathic concern (EC), and personal distress (PD). It provides four distinct scores, which makes it a multidimensional empathy measure. The PT subscale is considered to measure cognitive empathy, and the EC and PD subscales, affective empathy (Davis, 1983; De Corte et al., 2007). Finally, the F subscale does not appear to be easily categorized within the cognitive-affective conceptualization and may tap into both cognitive and affective aspects of empathy (Davis, 1980, 1983; De Corte et al., 2007).

Empathy is also often measured using behavioural tasks. An example is the Reading the Mind in the Eyes Test (RMET; Baron-Cohen, Wheelwright, Hill, Raste, & Plumb, 2001). This task consists of 36 pictures showing the eye region of an actor. Participants must choose the word that best describes the emotion felt by the person from among four words. Although commonly considered as a cognitive empathy task (Lawrence et al., 2004; Murphy & Lilienfeld, 2019), several authors argue that it is an emotion recognition task (Kittel, Olderbak, & Wilhelm, 2022; Oakley, Brewer, Bird, & Catmur, 2016). For example, a meta-analysis by Kittel et al. (2022) on the RMET's psychometric properties suggests that scores on the RMET are more strongly associated with scores on emotion perception tasks than those on theory of mind tasks or empathy questionnaires. Similarly, Oakley et al. (2016) found that autistic individuals had specific difficulties on a naturalistic theory of mind task, whereas individuals with alexithymia had difficulties on the RMET. This thus highlights one of the points of contention regarding empathy, namely whether this construct encompasses emotion recognition and if so, to what extent (Bollen, 2023). Nevertheless, to ensure that the current meta-analysis adequately represents the current state of empathy research in autism, the RMET and related tasks will be considered as assessing cognitive empathy.

A number of researchers also use more naturalistic empathy scenarios. In the empathic accuracy task, participants view videos in which someone (the target) recounts an autobiographical emotional event. While watching the video, they are asked to rate the target's emotional state (cognitive empathy measure), as well as their own emotional state (affective empathy measure; Zaki, Weber, Bolger, & Ochsner, 2009; but also see Ickes, Stinson, Bissonnette, & Garcia, 1990). There are also feigned distress tasks, which are often used with children. In these tasks, an examiner in the same room as the child or the child's parent usually pretends to hurt themself (Sigman, Kasari, Kwon, & Yirmiya, 1992; Zahn-Waxler, Radke-Yarrow, Wagner, & Chapman, 1992). The reactions and vocalizations of the child to the examiner's (or the parent's) simulated distress are filmed and subsequently coded.

Finally, certain tasks are combined with physiological and/or neuroimaging measures, typically fMRI. One such task is the self-other empathy task, in which participants are asked to rate both how they feel and how the other person feels in response to emotional stimuli, such as pictures or videos (Greimel et al., 2010; Regenbogen et al., 2012). There is also the empathy for pain task. In one variant of this task, participants are shown pictures or videos depicting either painful or non-painful situations (Avenanti, Minio-Paluello, Bufalari, & Aglioti, 2006; Lamm et al., 2011). In another variant, participants are typically placed in a prone position in a scanner and visual cues indicate whether they or another person in the room will receive a stimulation and whether it will be painful or not (Lamm et al., 2011; Singer et al., 2004). Participants are then often asked to rate the unpleasantness and/or intensity of the other person's pain and, in some cases, their own (Avenanti et al., 2006; Singer et al., 2004).

1.3. Empathy in autism

Empathy deficits contribute to the three socio-communicative criteria for an autism diagnosis in the fifth edition of the *Diagnostic and Statistical Manual of Mental Disorders* (DSM-5-TR; American Psychiatric Association, 2022). Decreased affect and emotion sharing is a part of criterion A1, *Deficits in socio-emotional reciprocity*. It is also a component of sign A2, *Deficits in nonverbal communicative behaviours used for social interaction*, in the form of deficits in understanding and using nonverbal communication, up to and including a total lack of facial expression or gestures. Finally, it can be integrated into sign A3, *Deficits in developing and maintaining relationships, appropriate to developmental level*, through the apparent absence of understanding or production of social context related to emotions. For example, maintaining a neutral face during a funeral can be scored in each of these three signs.

Similarly, dominant standardized diagnostic instruments, such as the observational scale Autism Diagnostic Observation Schedule (ADOS) and the retrospective Autism Diagnostic Interview-Revised (ADI-R), score socio-emotional difficulties as a sign of autism. For example, module 4 of the ADOS, which is used to assess autism in verbally fluent adolescents and adults, contains an item to score empathy and references to others' emotions (B6) (Lord et al., 2012). The Adult Asperger

Assessment (Baron-Cohen, Wheelwright, Robinson, & Woodbury-Smith, 2005) is another diagnostic instrument used to assess the verbal end of the autistic spectrum in adults, based on the DSM-IV-TR criteria. One of its criteria is a lack of socio-emotional reciprocity and it cites, as an example, a lack of empathy (Baron-Cohen, Wheelwright, et al., 2005). Moreover, it includes the Empathy Quotient (Baron-Cohen & Wheelwright, 2004) as a screening instrument (Baron-Cohen, Wheelwright, et al., 2005). This self-report questionnaire is also widely used in the scientific literature to measure "autistic traits". Scoring beyond the threshold of this instrument is taken as a demonstration of the presence of autistic traits in a large array of conditions, which raises multiple interpretation issues (Mottron & Bzdok, 2020).

Accordingly, the systematizing-empathizing model (Baron-Cohen, 2009, 2010; Wakabayashi et al., 2007), grounded on the sex-related differences noted in empathy, characterizes autistic people as hypo-empathizers. This model has deep connections with one of the purported neurobiological descriptions of autism, the extreme male brain theory (Baron-Cohen, 2002; Baron-Cohen, Knickmeyer, & Belmonte, 2005). Despite the major role of overt or covert manifestations of an empathy disorder in the diagnosis of autism, as well as in the resulting neurocognitive models constructed to account for them, current research on empathy in autism has provided contradictory results. Indeed, although some researchers have reported global empathy deficits in autism (e.g., Grove, Baillie, Allison, Baron-Cohen, & Hoekstra, 2014; Mathersul, McDonald, & Rushby, 2013), others have been unable to detect any empathy deficits (e.g., Bird et al., 2010; Newbigin, Uljarevic, Vivanti, & Dissanayake, 2016). In addition, a number of studies specifically found deficits in cognitive empathy (e.g., Dziobek et al., 2008; Fan, Chen, Chen, Decety, & Cheng, 2014) or suggested the presence of an atypical form of empathy (e.g., De Coster, Wiersema, Deschrijver, & Brass, 2018; Komeda, Kosaka, Fujioka, Jung, & Okazawa, 2019). This may suggest that the results are influenced by how empathy is operationalized, as there are several ways to assess it. Each measure uses a different definition and targets specific components of empathy. Clarifying the empathy profile in autism would be of great importance not only for revising the contribution of empathy competence in current models of autism, but also for distinguishing autism from other conditions conceptualized as empathy disorders, such as psychopathy and callous-unemotional traits.

1.4. Potential empathy moderators

Several factors can influence empathy, such as age, sex, and alexithymia. Empathy gradually develops over time (McDonald & Messinger, 2011). As early as 18 h after birth, newborns appear to be sensitive to other babies' distress. This is an expression of emotional contagion, which is a precursor to affective empathy (McDonald & Messinger, 2011; Singer, 2006). Personal distress then develops during infancy, which leads to the emergence of empathic concern and prosocial behaviour during the second year of life (Hoffman, 2000; McDonald & Messinger, 2011). Gains in cognitive empathy seem to appear later, around the age of 4 or 5 years, as the prefrontal cortex and temporal cortices further develop (Singer, 2006). At that age, children can generally take another person's perspective and start to understand another person's emotional experience (McDonald & Messinger, 2011). Empathy then continues to develop until early adulthood (McDonald & Messinger, 2011; Singer, 2006).

Sex and gender, which are deeply interconnected, can also influence empathy (Rochat, 2023). However, since studies typically categorize participants in a binary manner (i.e., male versus female) without distinguishing between sex and gender, the term "sex" was used. Females are, on average, more empathetic than males (Christov-Moore et al., 2014). Indeed, at around one to two years of age, girls show greater empathic concern than boys and engage more in comforting behaviours in response to other people's distress (Christov-Moore et al., 2014; Rochat, 2023). From two years of age, girls also perform better at false belief tasks and from nine years of age, they are, on average, better at understanding other people's intentions and feelings than boys (Christov-Moore et al., 2014). Upon reaching adolescence, the difference between males and females in terms of empathy appears to further increase (Christov-Moore et al., 2014; Rochat, 2023), with adolescent girls largely showing greater empathy than boys (Silke, Brady, Boylan, & Dolan, 2018), and this difference is maintained into adulthood (Christov-Moore et al., 2014; Rochat, 2023). The type of stimuli used and the context appear to impact empathy in males, but not in females, which suggests that, despite displaying less empathy than females in most situations, males have as much empathy as females in certain circumstances (Christov-Moore et al., 2014).

A final factor that has an impact on empathy is alexithymia (Bird & Viding, 2014). Alexithymia, as empathy and theory of mind, can be subdivided into a cognitive and an affective component. *Cognitive alexithymia* is characterized by a decreased ability to verbalize and recognize one's own emotions, whereas *affective alexithymia* is characterized by a decreased ability to fantasize and indicate whether an emotion arises from oneself or from an external stimulus (Bermond et al., 2007). A decreased capacity to reflect upon one's own emotions also appears to be common to both cognitive and affective alexithymia (Bermond et al., 2007). Alexithymia is associated with reduced empathy, as shown by several studies (e.g., Banzhaf et al., 2018; Bird et al., 2010; Grynberg, Luminet, Corneille, Grèzes, & Berthoz, 2010), and it often co-occurs with autism. Indeed, a recent meta-analysis found the prevalence of alexithymia within the autism spectrum to be 49 %, which was higher than in the neurotypical comparison group, for which the prevalence of alexithymia was 5 % (Kinnaird, Stewart, & Tchanturia, 2019). Furthermore, dissociating alexithymia from theory of mind is relevant for the interpretation of how autistic people perform in socio-communicative tasks (Rødgaard, Jensen, & Mottron, 2019).

1.4.1. Other possible moderators to consider in autism

Publication year is a factor that could specifically influence empathy in autism. Indeed, the meta-analysis of Rødgaard, Jensen, Vergnes, et al. (2019) showed that the difference between autistic and neurotypical individuals appeared to decline over the years for several psychological constructs, including emotion recognition and theory of mind. This finding is attributed to the changes in the diagnostic criteria for autism, which have become broader over time, thus increasing heterogeneity within the autism spectrum (Mottron & Bzdok, 2020; Rødgaard, Jensen, Vergnes, Soulieres, & Mottron, 2019).

The verbal intelligence quotient (VIQ) is another important factor to consider in autism, as autism is often associated with lower verbal abilities relative to non-verbal abilities (Audras-Torrent et al., 2021; Nader, Jelenic, & Soulieres, 2015). Thus, autistic individuals may have more difficulties than typical individuals to complete empathy measurement instruments that rely heavily on verbal abilities, despite having a full-scale IQ in the normal range. Moreover, although alexithymia seems to influence the responses of participants in terms of physiological and neuroimaging measures, verbal IQ appears to have a stronger impact than alexithymia on their performance on behavioural measurement instruments (Sivathasan, Fernandes, Burack, & Quintin, 2020). Thus, alexithymia and verbal IQ may differentially influence empathy.

1.5. Previous reviews on empathy in autism

Several systematic reviews on empathy in autism have been published over the last decade. One of those investigated cognitive, affective, and motor empathy in children and adolescents with autism or conduct disorder, with a focus on studies using physiological empathy measures (Bons et al., 2013). It found mixed results for cognitive empathy and lower or delayed motor empathy in autistic children and adolescents. However, no studies assessing affective empathy as an autonomic response to emotional stimuli were found in autism, preventing conclusions from being drawn (Bons et al., 2013). Another review covered self-reported empathy in adult autistic women (Kok, Groen, Becke, Fuermaier, & Tucha, 2016). This review identified six studies comparing autistic women and men, all of which used the EQ. It showed that autistic women scored similarly to autistic men on the EQ, whereas typical women reported higher scores than typical men. Autistic women also had lower EQ scores than typical women and, to a lesser extent, than typical men (Kok et al., 2016). A third review explored the influence of sex, age, IQ, and co-occurring conditions on empathy in autism (Harmsen, 2019). It found that typical sex differences in empathy are reduced in autism and that empathic abilities tend to improve with age. It also suggested that language skills and executive functioning may be more closely linked to empathy than IQ alone. Co-occurring ADHD may exacerbate empathy difficulties, while psychosis may have a mitigating effect. Finally, the review examined empathy-related diagnostic tools and interventions, as well as the putative genetic, neural and sensory mechanisms underlying the empathy deficits reported in autism (Harmsen, 2019).

A meta-analysis investigated both trait and state cognitive empathy, empathic concern, and empathic accuracy in autistic children and adults, as well as the effect of culture, gender, and age (Song, Nie, Shi, Zhao, & Yang, 2019). The authors found lower trait and state cognitive empathy and empathic concern, similar state empathic accuracy, and greater trait empathic accuracy in the autistic group compared to the typical group. They further showed that culture did not moderate empathy, that gender moderated trait empathic concern and empathic accuracy, and state cognitive empathy, and that age moderated trait cognitive empathy, state empathic concern, and both trait and state empathic accuracy (Song et al., 2019). Finally, another meta-analysis reviewed 35 studies concomitantly measuring both cognitive and affective empathy in autistic children and adults and assessed the impact of age, culture, and the type of measure used (questionnaire versus picture or video based) as moderators (Fatima & Babu, 2023). It found that the autistic group had lower cognitive and affective empathy than the typical group, with large and moderate effect sizes, respectively. It further showed that age and culture did not moderate empathy, whereas the type of measure moderated cognitive empathy (Fatima & Babu, 2023).

1.6. The present systematic review and meta-analysis

Although several previous systematic reviews examined empathy in autism, what is unique to this systematic review and meta-analysis is its focus on the empathy measures that were used. The first objective was to assess whether autistic individuals have an empathy deficit and, if so, to identify in which empathy component it lies. Thus, we investigated cognitive and affective empathy by comprehensively synthesizing research on empathy in autism. In addition, we assessed the impact of several variables on empathy, namely, age and sex, but also alexithymia, verbal IQ, publication year, and study quality, which were not evaluated in previous meta-analyses. The second objective was to determine whether the measure used to quantify empathy had an impact on the results of the studies. Thus, the three most frequently used empathy measures were separately analyzed to assess whether different measures led to different results concerning empathy in autism. Finally, the inclusion of a qualitative literature review allowed for the consideration of results from naturalistic empathy tasks or neuroimaging tasks that were often excluded from previous reviews.

2. Methods

This literature review and meta-analysis was conducted according to the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) 2020 statement (Page et al., 2021) and was pre-registered in Prospero (registration number: CRD42020205343). All data and analysis codes are available on Open Science Framework (https://osf.io/hp7as/).

2.1. Eligibility criteria

All empirical studies quantifying empathy in autism were considered. The inclusion criteria were the following: 1) the article had to compare an autistic group with a non-clinical control group (henceforth called the typical group), 2) it had to be written in English or French, 3) it had to be published in a peer-reviewed journal, and 4) it had to use a quantitative empathy measure (see section 2.2. for measure selection). Papers that 1) did not report new data (e.g., reviews) or 2) described interventions to improve the empathy of autistic individuals were excluded. Furthermore, papers were excluded from the review if insufficient data were provided to calculate an effect size. When several studies used an overlapping sample, only the study with the largest number of participants was included. Finally, studies reporting data that violated the normality assumption were only included in supplemental analyses, since it may prevent from calculating a reliable effect size (Sun & Cheung, 2020). In addition, studies using more naturalistic empathy measures or neuroimaging tasks were qualitatively synthesized as they can provide additional insights regarding empathy in autism.

2.2. Search strategy

A two-step search strategy was used for this meta-analysis. The first search aimed to clarify the exact keywords to be used by identifying empathy tasks. A second search was then conducted by adding these tasks as search terms. The first search was conducted in the PsycNET and PubMed databases from July to August 9, 2019. The search terms were empath* AND (autis* OR asperger*). After excluding duplicates, 1268 records of interest remained. After screening the titles and abstracts of the articles, the full text was retrieved for articles deemed to be of interest to ensure they fulfilled the eligibility criteria. In total, 109 articles were retained following the initial search. However, there are conflicting conceptualizations of empathy: the same task can be judged by some authors to be an empathy task, but not by others. As such, to ensure that all relevant articles were considered even if they did not mention empathy, an additional step was taken. The description of nine empathy models was analyzed and the words used to define empathy throughout the text in each were extracted. These words were then classified based on their frequency of occurrence across the empathy models. Finally, the tasks deemed to measure empathy in the 109 articles were listed. If three words appeared at least twice in the task description of the original article, it was considered to be an empathy task and was included as a search term for the final search. This preliminary step was necessary to ensure that the meta-analysis was exhaustive and identified articles which used an empathy task but did not mention the concept. The references for the selected empathy models are provided in Supplementary Materials S1.

A second literature search was conducted on September 21, 2020, and was updated on January 7, 2022, May 1st, 2024, and January 17, 2025. Studies were identified by searching the PsycNET, PubMed, Scopus, and Academic Search Complete electronic databases. The search terms were (empath* OR EMPATHY_TASK_NAME) AND (autis* OR asperger*) and the search was performed across all fields. The full search strings and filters used for each database can be found in Supplementary Materials S2. This two-step procedure made it possible to include all studies using an empathy measure, independently of their focus. In addition, the reference list of previous systematic literature reviews and meta-analyses on empathy in autism were screened to identify additional studies (Bollen, 2023; Fatima & Babu, 2023; Harmsen, 2019; Kok et al., 2016; Song et al., 2019).

2.3. Study selection

The identified records were exported to EndNote X9 software. After excluding the duplicates following the literature search in the databases, 3070 papers of interest remained. The titles and abstracts were then screened for relevance by one author. If the eligibility criteria were met or insufficient information was available in the abstract, the full text was retrieved for further assessment of the eligibility criteria. If the eligibility was unclear even after reading the full-text article, the article was discussed with a second reviewer to reach a consensus. In total, 179 articles were excluded after reading the full text. The references for these articles are provided in Supplementary Materials S3. Finally, 226 articles were retained for the review: 205 were included in the meta-analysis and 37 in the literature review (see Supplementary Materials S4 and S5 for references). No additional articles were included following the reference search from previous systematic literature reviews and meta-analyses (see Fig. 1).

2.4. Data extraction

The following data were extracted from every article that met the eligibility criteria: article information (author, year, and title), sample characteristics (sample size, age, sex, IQ, diagnosis, mean alexithymia score), name and characteristics of the empathy measure(s), and the main conclusions of the study. The data required to calculate the effect size for each available empathy measure within an article (the mean and standard deviation for each group, a *t*-value, or an *F* value from a one-way ANOVA) were also retrieved from every article included in the meta-analysis. When relevant data was missing, authors were contacted by email. A modified version of Polanin and Terzian (2019) data-sharing agreement form was used to explain why the data was requested and under which conditions it would be used.

Data were first extracted by one reviewer and one trained research assistant. A second reviewer then independently collected the data for 53 articles (23 %). Percent agreement between the two raters was of 93.14 %. In cases of discrepancy, a discussion took place to reach consensus.

2.5. Data preparation process

A number of articles reported the mean and standard error instead of the standard deviation, in which case the standard error was converted to a standard deviation using the formula described in the *Cochrane Handbook for Systematic Reviews of Interventions* (Higgins et al., 2022). Furthermore, in certain cases, separate scores were given for two subgroups of the population of interest (e.g., separate scores were given for men and women). The scores of the subgroups were thus combined according to the method recommended by the *Cochrane Handbook for Systematic Reviews of Interventions* (Higgins et al., 2022). Finally, when

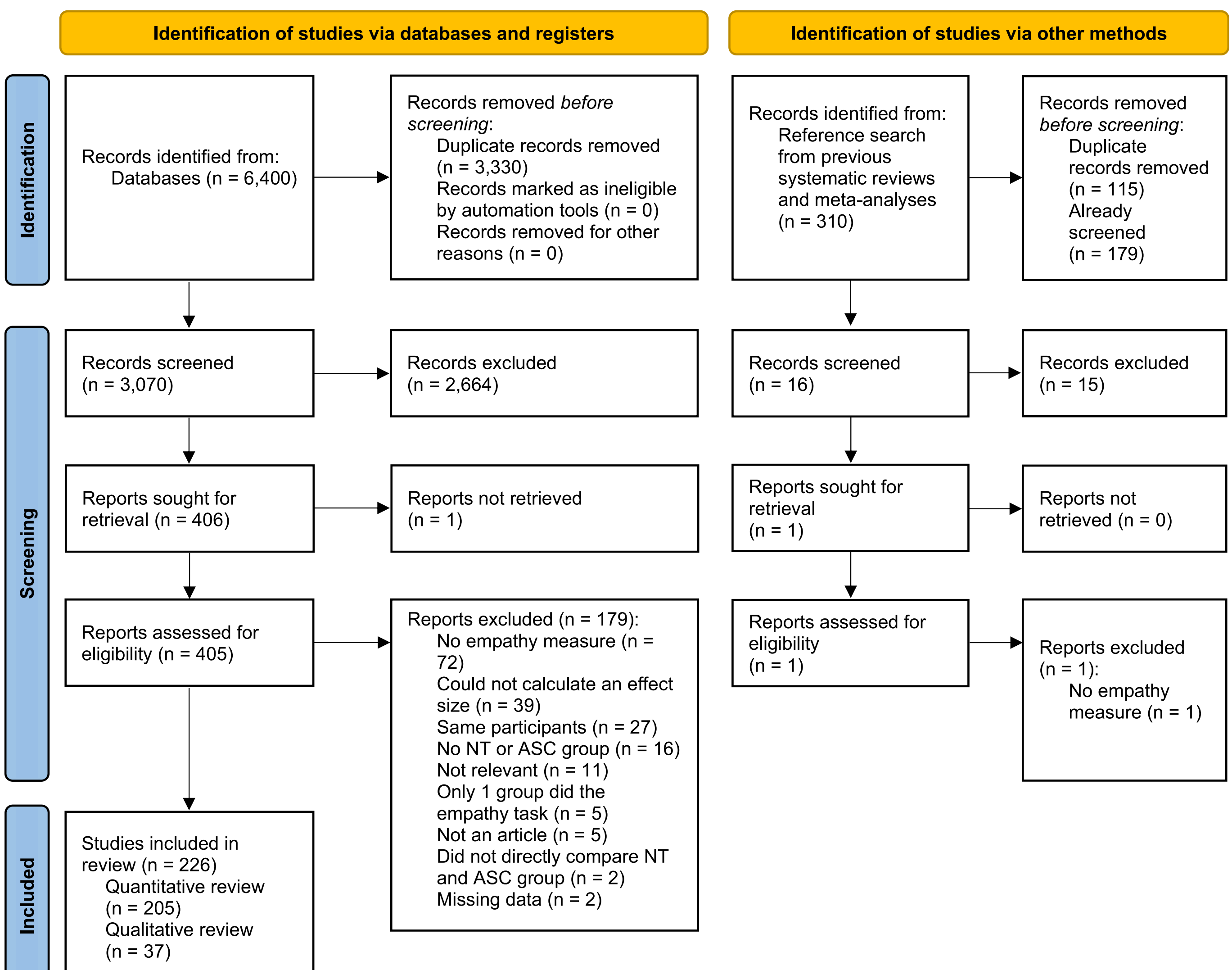

Fig. 1. PRISMA flow diagram. PRISMA diagram showing the study selection process and the final number of studies included in the quantitative review (i.e., meta-analysis) and the qualitative review.

necessary, the results of one-way ANOVAs between the autistic and typical groups were converted into a *t*-test using the following formula: $F = t^2$. As the *F* statistic is strictly positive, this manipulation was sometimes required to obtain an effect size with the appropriate sign.

2.6. Quality assessment

The quality of individual studies was assessed using an adapted version of the Newcastle-Ottawa Scale (NOS), such as that used by Rødgaard, Jensen, Vergnes, et al. (2019). This tool comprises eight items. Four items assess the selection process of the participants, one determines comparability between the autistic and typical groups, and three evaluate the study's outcome. A maximum score of 9 can be achieved for high-quality studies. Study quality was first assessed by one reviewer and one trained research assistant. A second reviewer then independently applied the NOS to 53 articles (23 %). Inter-rater agreement was of 80.42 %. Any discrepancy led to a discussion between the two reviewers to reach consensus.

2.7. Statistical analyses

All statistical analyses were conducted using RStudio, version 4.1.2 (R Core Team, 2021). The autistic and typical group scores for different empathy measures were compared by computing the standardized mean differences, more specifically Hedges' g, using the *esc* package (version 0.5.1; Lüdecke, 2019). Positive standardized mean differences indicate that the autistic group had greater empathy scores than the typical group, whereas negative standardized mean differences indicate that the typical group had greater empathy scores. Following Cohen's convention, Hedges' g is small if $g \approx 0.2$, medium if $g \approx 0.5$, and large if $g \approx 0.8$ (Cohen, 1988).

The meta-analyses were performed using the *metafor* package (version 4.6.0; Viechtbauer, 2010). First, meta-analyses were conducted to synthesize the study results concerning cognitive, affective, and unidimensional empathy. Then, separate meta-analyses were performed for the three most frequently used empathy measures to assess their individual impact on the study results. The empathy measures were the EQ, the RMET, and the IRI. When both adult and child versions of the measure were available, distinct meta-analyses were conducted.

Either a random-effects model or a three-level model was applied. Since the random-effects model postulates effect size independence (Cheung, 2019), whenever several effect sizes deriving from the same article contributed to the same analysis, a three-level model was used. More specifically, for unidimensional empathy and the most highly used empathy measures, a random-effects model was applied, for which weights were assigned using the inverse-variance method. As the random-effects model accounts for the presence of heterogeneity across studies, it is recommended when both within-study estimation error variance and between-study variance are expected (Borenstein, Hedges, Higgins, & Rothstein, 2010). Furthermore, this method makes it possible to account for smaller studies, as it assumes that all studies report a different effect size rather than their having a similar true effect size (Borenstein et al., 2010). For cognitive and affective empathy, as several studies reported multiple effect sizes, a three-level model was used (Van den Noortgate, López-López, Marín-Martínez, & Sánchez-Meca, 2013, 2015). The first level represented the effect size for every empathy task within each study, the second level the pooled effect size for each study, and the third level the summary effect size for all included studies. Heterogeneity variance (τ^2) was calculated using the restricted maximum likelihood estimator (REML). Furthermore, the confidence intervals for the summary effects were calculated using the Knapp and Hartung adjustment (Hartung & Knapp, 2001) when a random-effects model was applied, as previously suggested (Langan et al., 2019), whereas it was calculated using a *t*-distribution when a three-level model was applied. Influential cases were identified after each meta-analysis following the procedure and thresholds recommended by Viechtbauer and Cheung (2010). The meta-analyses were rerun without the influential cases when they were identified.

To assess the impact of several moderators on cognitive, affective, and unidimensional empathy, meta-regressions were conducted with a mixed-effects model or a three-level mixed-effects model. Specifically, the moderators were publication year, study quality (assessed using the NOS), the standardized mean difference (Hedge's g) for alexithymia scores (measured using the Toronto Alexithymia Scale), the standardized mean difference (Hedge's g) for verbal IQ, age group (children or adult), and sex (male or female). As before, the REML estimator and Knapp and Hartung adjustment (Knapp & Hartung, 2003) or a *t*-distribution were used. All meta-regressions were performed excluding influential cases. Additionally, to directly compare the effect size obtained for cognitive and affective empathy, a three-level mixed-effects model, with the empathy component assessed (cognitive or affective) set as a moderator, was done. Exploratory sensitivity analyses restricted to high-quality studies (as defined by a score of 7 or over on the NOS) were performed to assess the robustness of the results for cognitive, affective, and unidimensional empathy, as well as for the most highly used empathy measures. Similarly, supplemental analyses were also performed by including studies with non-normal data, since such studies are usually included in meta-analyses. Finally, the average coefficient of variation across studies was calculated for both the autistic and typical groups to compare the relative variability in their empathy scores.

2.7.1. Heterogeneity across studies

Heterogeneity across studies was tested by computing the τ^2 and I^2 statistics. τ^2 is an indicator of between-study variance (Borenstein et al., 2010). Similarly, I^2 assesses the variability across studies that is caused by heterogeneity rather than by chance (Deeks, Higgins, & Altman, 2019; Higgins, Thompson, Deeks, & Altman, 2003). I^2 values of 25 %, 50 %, and 75 % indicate a low, moderate, and high percentage of heterogeneity, respectively, across studies (Higgins et al., 2003). Forest plots were generated and inspected to further assess the presence of heterogeneity across studies (Phan, Tian, Cao, Black, & Yan, 2015). Prediction intervals were also calculated, as they represent the expected range of effects for future studies based on current study heterogeneity (IntHout, Ioannidis, Rovers, & Goeman, 2016).

In addition to the *metafor* package (Viechtbauer, 2010), the var.comp function of the *dmetar* package (version 0.0.9000; Harrer, Cuijpers, Furukawa, & Ebert, 2019) was used to calculate the I^2 statistic for the three-level model. Furthermore, the code provided by Fernández-Castilla et al. (2020) was used to visualize the forest plots for the three-level meta-analyses.

2.7.2. Risk of bias across studies

The risk of bias across studies was first assessed through visual inspection of funnel plots by one reviewer. The funnel plots were obtained by plotting the standardized mean difference for the individual studies against the standard error of the studies (Egger, Smith, Schneider, & Minder, 1997). The trim-and-fill method of Duval and Tweedie was then applied (Duval & Tweedie, 2000a). This method makes it possible to detect and correct for publication bias by estimating the number of missing studies and filling them in in the funnel plot (Shi & Lin, 2019). Both L_0 and R_0 estimates were calculated, as they show better mean square error properties than Q_0 (Duval & Tweedie, 2000b). Finally, the three-parameter selection model (McShane, Bockenholt, & Hansen, 2016; Pustejovsky & Rodgers, 2019) was used to assess the impact of publication bias by modelling two hypothetical publication processes. The first three-parameter selection model assessed whether studies with statistically significant results (as indicated by a *p*-value of 0.05 or less) were more likely to be published than studies with non-significant results. The second three-parameter selection model assessed whether studies with either significant or marginally significant results (as indicated by a *p*-value of 0.10 or less) were more likely to be published than truly non-significant results.

As the trim-and-fill method and the three-parameter selection model assume that effect sizes are independent, a modified version of Egger's regression test was used to assess the risk of bias across studies for the three-level meta-analyses (i.e., for cognitive and affective empathy). Specifically, it integrates a multilevel meta-analysis to account for the dependence of effect sizes (Rodgers & Pustejovsky, 2021) and uses $\sqrt{\frac{n_1+n_2}{n_1 n_2}}$ as a predictor, rather than the classical standard error formula, to avoid inflating Type I error (Pustejovsky & Rodgers, 2019). Two different funnel plots were also generated following each three-level meta-analysis using the code provided by Fernández-Castilla et al. (2020). First, the standardized mean difference for the individual studies was plotted against the standard error of the studies and then the averaged standardized mean difference for the studies was plotted against the meta-analytic standard error, as suggested by Fernández-Castilla et al. (2020).

3. Results

3.1. Study characteristics

A summary of the sample characteristics can be found in Tables S1 and S2. In total, 13,278 autistic participants and 44,218 typical participants were included in the meta-analysis. The most often reported empathy values came from the Reading the Mind in the Eyes Test (RMET; cognitive empathy; used in 25.30 % of studies), the Empathy Quotient (EQ; unidimensional empathy; used in 20 % of studies), and the Interpersonal Reactivity Index (IRI; multidimensional empathy; used in 19.50 % of studies). See Table S3 in the supplementary materials for more details concerning which studies were included in each analysis. The summary statistics and effect sizes for each study included in the cognitive, affective, and unidimensional empathy analyses can be found in Tables S4-S6 in the supplementary materials.

In addition, the study results for more naturalistic empathy tasks or neuroimaging tasks were reviewed in the qualitative literature review, namely the feigned distress task ($n = 16$), the empathic accuracy task ($n = 4$), the self-other empathy task ($n = 5$), and the empathy for physical pain task ($n = 12$).

3.2. Quality assessment

In terms of the quality of individual studies, 34 studies obtained a score between 3 and 5 on the Newcastle-Ottawa Scale. Another 52 studies received a moderate score (between 5 and 7) and 103 studies, a high score (7 and over). Finally, one study received a moderate score for one task and a high score for the other due to the loss of participants for one of the tasks. Thus, the quality of the studies included in the meta-analysis was relatively good, with about half of the studies obtaining a high score. These were the studies subsequently included in the sensitivity analyses for high-quality studies. For more details concerning the quality of each of the included studies, see Tables S7 and S8.

3.3. Cognitive, affective, and unidimensional empathy

A summary of the results obtained for cognitive, affective, and unidimensional empathy is shown in Table 1.

3.3.1. Cognitive empathy

The three-level model comparing cognitive empathy scores between autistic and typical participants ($k = 178$) revealed a large effect size of $g = -0.85$ (95 % CI [−0.93; −0.76], $p < .001$), indicating that the autistic group had significantly lower cognitive empathy scores than the typical group (see Fig. S1). $I^2_{\text{Level 2}}$, which represents the amount of heterogeneity variance within studies, accounted for 44.94 % of variance, whereas between-study heterogeneity ($I^2_{\text{Level 3}}$) accounted for 30.46 % of the total variance. This indicates that there was more heterogeneity within studies than between different studies. Rerunning analyses without one influential case ($k = 177$) showed similar results ($g = -0.83$, 95 % CI [−0.91; −0.76], $p < .001$; Fig. S2). $I^2_{\text{Level 2}}$ and $I^2_{\text{Level 3}}$ now accounted for 44.59 % and 29.30 %, respectively, of the total variance.

3.3.2. Affective empathy

The three-level model comparing affective empathy scores between autistic and typical participants ($k = 107$) revealed a small effect size of $g = -0.17$ (95 % CI [−0.31; −0.03], $p = .02$), indicating that the autistic group had slightly lower affective empathy scores than the typical group (see Fig. S3). $I^2_{\text{Level 2}}$ and $I^2_{\text{Level 3}}$ accounted for 89.52 % and 0 %, respectively, of the total variance, suggesting that almost all the variance is explained by level 2 (i.e., differences within studies). A direct comparison of the cognitive and affective empathy effect sizes further showed that the difference in cognitive empathy between autistic and typical participants was significantly greater than the difference in affective empathy ($F(1, 282) = 80.95$, $p < .0001$).

3.3.3. Unidimensional empathy

The random-effects model comparing unidimensional empathy between autistic and typical participants ($k = 64$) revealed a large effect size of $g = -1.70$ (95 % CI [−1.85; −1.56], $p < .001$; Fig. S4). This suggests that the autistic group had significantly lower unidimensional empathy scores than the typical group. However, heterogeneity across studies was very high, as indicated by an I^2 value of 94.49 %.

3.3.4. Publication bias

Inspection of the funnel plots for cognitive empathy suggests a departure from symmetry, which might indicate the presence of publication bias, whereas inspection of the funnel plots for affective empathy

Table 1
Meta-analysis results for cognitive, affective and unidimensional empathy.

Empathy component	*n*	*k*	*g*	*SE*	*t* value	*df*	*p* value	95 % CI	95 % PI	τ^2	τ^2 L2	τ^2 L3	I^2	I^2 L2	I^2 L3
Cognitive empathy															
Cognitive	137	178	−0.85	0.04	−20.50	177	<0.001	[−0.93; −0.76]	[−1.70; 0.01]		0.11	0.08		44.94	30.46
Cognitive no outliers	136	177	−0.83	0.04	−20.83	176	<0.001	[−0.91; −0.76]	[−1.66; −0.01]		0.10	0.07		44.59	29.30
Affective empathy															
Affective	62	107	−0.17	0.07	−2.35	106	0.020	[−0.31; −0.03]	[−1.54; 1.20]		0.47	0.00		89.52	0
Unidimensional empathy															
Unidimensional	64	64	−1.70	0.07	−23.08	63	<0.001	[−1.85; −1.56]	[−2.75; −0.65]	0.27			94.49		

n = number of studies; *k* = number of effect sizes; *g* = Hedges' g; 95 % PI = 95 % prediction intervals; L2 = Level 2 (heterogeneity within studies); L3 = Level 3 (heterogeneity between studies).

indicates only slight asymmetry (see Fig. S5). The results of Egger's regression test suggest the presence of small-study effects for cognitive empathy ($t(175) = -4.98$, $p < .0001$) but not for affective empathy ($t(105) = -0.73$, $p = .47$).

Inspection of the funnel plot for unidimensional empathy suggests a departure from symmetry, potentially indicating the presence of publication bias (see Fig. S9). Concerning the trim-and-fill analysis, the R_0 estimator did not impute any missing studies, but the L_0 estimator suggested that seven studies were missing on the right (see Table S9). Finally, the three-parameter selection model suggests that studies with statistically significant or marginally significant results were as likely to be published as studies with truly non-significant results (see Table S10). Thus, the three-parameter selection model did not indicate the presence of significant publication bias. Furthermore, the estimated effect sizes were within the same range as those obtained from the random-effects meta-analysis.

3.3.5. Meta-regressions

Publication year was not a significant moderator for either cognitive ($B = 0.01$, $p = .11$, $k = 177$), affective ($B = 0.01$, $p = .56$, $k = 107$), or unidimensional empathy ($B = 0.00$, $p = .76$, $k = 64$). Similarly, study quality was not a significant moderator for either cognitive ($B = 0.02$, $p = .51$, $k = 177$), affective ($B = 0.04$, $p = .37$, $k = 107$), or unidimensional empathy ($B = 0.03$, $p = .43$, $k = 64$). This indicates that neither publication year nor study quality had a significant impact on the difference in empathy scores between autistic and typical participants.

Alexithymia could only be tested as a moderator for cognitive and affective empathy. It was not a significant moderator for either cognitive ($B = -0.09$, $p = .59$, $k = 25$) or affective empathy ($B = -0.07$, $p = .67$, $k = 29$). Similarly, verbal IQ was not a significant moderator for either cognitive ($B = -0.31$, $p = .10$, $k = 65$), affective ($B = -0.37$, $p = .28$, $k = 27$), or unidimensional empathy ($B = -0.03$, $p = .97$, $k = 14$). Thus, neither alexithymia nor verbal IQ had an influence on the difference between the autistic and typical participants empathy scores.

The age group of the participants was not a significant moderator for either cognitive ($B = 0.02$, $p = .87$, $k = 114$), affective ($B = 0.31$, $p = .07$, $k = 74$), or unidimensional empathy ($B = 0.04$, $p = .86$, $k = 43$). This indicates that the difference in empathy scores between autistic and typical children is not significantly different from that between autistic and typical adults.

Finally, sex could only be tested as a moderator for cognitive and unidimensional empathy. Sex was not a significant moderator for cognitive empathy ($B = -0.04$, $p = .87$, $k = 35$) but it was for unidimensional empathy ($B = -0.44$, $p = .01$, $k = 52$). Specifically, the results suggest that, relative to typical women, autistic women had more difficulties on unidimensional empathy tasks than autistic men relative to typical men. For more details concerning the results of the meta-regressions, see Table S11.

3.4. Empathy measures

A summary of the results obtained for the most frequently used empathy measures are presented in Table 2. The forest plots for each empathy measure are shown in Figs. S6 to S8 in the supplementary materials.

3.4.1. The Reading the Mind in the Eyes Test (RMET) and the Reading the Mind in the Eyes Test – Child (RMET-C)

For the RMET, a large effect size of $g = -0.76$ ($k = 48$, 95 % CI [−0.90; −0.61], $p < .001$) was found, indicating that autistic adults had significantly lower RMET scores than typical adults. The I^2 value was 65.87 %, which suggests moderate to high heterogeneity across studies. Rerunning the analyses without one influential case ($k = 47$) yielded similar results ($g = -0.72$, 95 % CI [−0.83; −0.60], $p < .001$). However, the heterogeneity across studies substantially decreased, with an I^2 value of 51.37 %. This indicates moderate heterogeneity across studies.

For the RMET-C, a large effect size of $g = -0.71$ ($k = 21$, 95 % CI [−0.89; −0.53], $p < .001$) was obtained, indicating that autistic children had significantly lower scores on the RMET-C than typical children. The I^2 value was 51.16 %, which indicates moderate heterogeneity across studies.

3.4.2. The Empathy Quotient (EQ) and Empathy Quotient-Child (EQ-C)

A large effect size of $g = -1.70$ ($k = 38$, 95 % CI [−1.87; −1.53], $p < .001$) was obtained for the EQ. This indicates that autistic adults had significantly lower scores on the EQ than typical adults. The I^2 value was 93.99 %, suggesting very high heterogeneity across studies.

For the EQ-C, a large effect size of $g = -1.84$ ($k = 8$, 95 % CI [−2.38; −1.31], $p < .001$) was obtained, which indicates that autistic children had lower empathy scores on this measure than typical children. However, the heterogeneity across studies was very high ($I^2 = 91.73$ %). Results for the EQ-C should be interpreted with caution due to the small number of studies included.

3.4.3. The Interpersonal Reactivity Index (IRI)

Separate meta-analyses were performed for each subscale of the IRI, namely perspective taking (PT), fantasy (F), empathic concern (EC), and personal distress (PD). The random-effects model comparing PT scores between autistic and typical participants revealed a large effect size of $g = -0.99$ ($k = 36$, 95 % CI [−1.11; −0.87], $p < .001$). This indicates that

Table 2
Meta-analysis results for the most frequently used empathy measures.

Empathy measure	k	g	SE	t value	df	p value	95 % CI	95 % PI	τ^2	95 % CI	I^2	95 % CI
RMET												
RMET	48	−0.76	0.07	−10.49	47	<0.001	[−0.90; −0.61]	[−1.45; −0.06]	0.11	[0.08; 0.37]	65.87	[58.59; 86.19]
RMET no outliers	47	−0.72	0.06	−12.33	46	<0.001	[−0.83; −0.60]	[−1.23; −0.20]	0.06	[0.03; 0.19]	51.37	[33.14; 76.81]
RMET-C	21	−0.71	0.09	−8.28	20	<0.001	[−0.89; −0.53]	[−1.31; −0.11]	0.07	[0.01; 0.26]	51.16	[16.74; 78.25]
EQ												
EQ	38	−1.70	0.08	−20.26	37	<0.001	[−1.87; −1.53]	[−2.60; −0.79]	0.19	[0.11; 0.39]	93.99	[89.81; 96.91]
EQ-C	8	−1.84	0.23	−8.09	7	<0.001	[−2.38; −1.31]	[−3.38; −0.31]	0.37	[0.14; 1.68]	91.73	[80.27; 98.05]
IRI												
IRI-PT	36	−0.99	0.06	−16.18	35	<0.001	[−1.11; −0.87]	[−1.43; −0.55]	0.04	[0.01; 0.16]	38.73	[9.37; 70.46]
IRI-PT no outliers	35	−1.02	0.06	−17.75	34	<0.001	[−1.13; −0.90]	[−1.34; −0.70]	0.02	[0.00; 0.15]	23.46	[0.00; 67.57]
IRI-F	25	−0.42	0.07	−5.76	24	<0.001	[−0.57; −0.27]	[−0.73; −0.11]	0.02	[0.00; 0.27]	19.09	[0.00; 78.88]
IRI-F no outliers	24	−0.43	0.08	−5.18	23	<0.001	[−0.60; −0.26]	[−0.86; 0.01]	0.04	[0.00; 0.29]	30.27	[0.00; 77.19]
IRI-EC	36	−0.59	0.07	−8.72	35	<0.001	[−0.73; −0.45]	[−1.15; −0.03]	0.07	[0.03; 0.22]	52.56	[28.69; 77.13]
IRI-PD	31	0.67	0.13	5.37	30	<0.001	[0.42; 0.93]	[−0.54; 1.89]	0.34	[0.20; 0.82]	83.36	[74.59; 92.40]

k = number of effect sizes; g = Hedges' g; 95 % PI = 95 % prediction intervals; RMET = Reading the Mind in the Eyes Test – revised version; RMET-C = Reading the Mind in the Eyes Test – Child version; EQ = Empathy Quotient; EQ-C = Empathy Quotient – Child; IRI-PT = Interpersonal Reactivity Index – Perspective Taking subscale; IRI-F = Interpersonal Reactivity Index – Fantasy subscale; IRI-EC = Interpersonal Reactivity Index – Empathic Concern subscale; IRI-PD = Interpersonal Reactivity Index – Personal Distress subscale.

the autistic group had significantly lower PT scores than the typical group. The heterogeneity across studies was low to moderate (I^2 = 38.73 %). Rerunning the analyses without one influential case (k = 35) showed similar results (g = −1.02, 95 % CI [−1.13; −0.90], p < .001). Moreover, the heterogeneity across studies decreased, with an I^2 value of 23.46 %. For the F subscale, a moderate effect size of g = −0.42 (k = 25, 95 % CI [−0.57; −0.27], p < .001) was obtained, indicating that the autistic group had significantly lower F scores than the typical group. The I^2 value was 19.09 %, which indicates low heterogeneity across studies. Similar results were obtained when rerunning the analyses without one influential case (k = 24, g = −0.43, 95 % CI [−0.60; −0.26], p < .001), although heterogeneity across studies increased, with an I^2 value of 30.27 %.

Concerning the EC subscale, a moderate effect size of g = −0.59 (k = 36, 95 % CI [−0.73; −0.45], p < .001) was found. This indicates that the autistic group had significantly lower EC scores than the typical group. The I^2 value was 52.56 %, which indicates moderate heterogeneity across studies. Finally, a moderate to large effect size of g = 0.67 (k = 31, 95 % CI [0.42; 0.93], p < .001) was obtained for the PD subscale, indicating that the autistic group had significantly higher PD scores than the typical group. The I^2 value was 83.36 %, which indicates high heterogeneity across studies.

3.4.4. Publication bias

Visual inspection of the funnel plots did not indicate asymmetry for the RMET, RMET-C, or EQ, but showed asymmetry for the EQ-C, suggesting possible publication bias for this measure (see Figs. S9 and S10). Concerning the trim-and-fill analyses, no studies were imputed for the RMET-C. In addition, the R_0 estimator only imputed one study to the left for the EQ whereas the L_0 estimator imputed one study to the left for the EQ-C. For the RMET, however, although the R_0 estimator only imputed two studies on the right, the L_0 estimator suggested that eight studies were missing on the right. For the IRI, the funnel plots showed no asymmetry for the PT, F, or EC subscales, but there was a slight departure from symmetry for the PD subscale. The trim-and-fill analysis with the R_0 estimator imputed two studies to the right for the F subscale and four studies to the right for the EC subscale, whereas the analysis with the L_0 estimator imputed nine studies to the left for the PD subscale. The estimated effect sizes, nevertheless, remained within the same range for all empathy measures once the imputed studies were considered, except for the PD subscale (see Table S9).

The three-parameter selection model did not suggest the presence of publication bias due to the publication process for the RMET, the EQ, and for the four subscales of the IRI (i.e., PT, F, EC, and PD). Thus, studies with statistically or marginally significant results were as likely to be published as studies with clearly non-significant results (see Table S10). Once again, the estimates were within the same range as those obtained from the random-effects meta-analysis.

Overall, these results suggest that there was little evidence for publication bias for the RMET-C, EQ, or EQ-C. Similarly, both the trim-and-fill method and the three-parameter selection model suggest the absence of substantial publication bias for three of the IRI subscales (i.e., PT, F, and EC). However, these analyses suggest the presence of publication bias for the RMET and for the PD subscale of the IRI.

3.5. Sensitivity analyses

We performed sensitivity analyses to assess the robustness of the results by including only high-quality studies (n = 104). The results remained within the same range for cognitive and unidimensional empathy, as well as for the RMET, RMET-C, EQ, and IRI. However, the effect size for affective empathy became non-significant (g = −0.11, 95 % CI [−0.29; 0.08], p = .25). Furthermore, the sensitivity analysis for the EQ-C could not be conducted, as only three studies remained. Supplemental analyses including studies with non-normal data were also performed. Once again, the results remained within the same range for cognitive, affective, and unidimensional empathy, as well as for the different empathy measures. For more details, see Tables S12 to S15 in the supplementary materials.

3.6. Coefficient of variation comparison analysis

After excluding aberrant values (n = 8), the average coefficient of variation across studies was of 35.06 % for the autistic group and of 22.64 % for the typical group. This suggests that the autistic participants exhibited greater relative variability in empathy scores compared to the typical participants. For additional details, see Supplementary Materials S6.

3.7. Qualitative systematic review

3.7.1. Feigned distress task

Several studies investigated empathy in autistic children and adolescents by assessing their responsiveness to distress. Most studies in which either the experimenter or the child's parent feigned distress found that autistic children and adolescents showed less concern and seemed less distressed than typical children and adolescents (Butean, Costescu, & Dobrean, 2014; Campbell, Moore, Northrup, & Brownell, 2017; Charman et al., 1997; Hutman et al., 2010; Li, Blijd-Hoogewys, Stockmann, Vergari, & Rieffe, 2023; McDonald & Messinger, 2012; Paz et al., 2024; Rieffe et al., 2021; Scambler, Hepburn, Rutherford, Wehner, & Rogers, 2007). However, two studies found that autistic children and adolescents showed as much concern as typical children and adolescents in response to the experimenter's distress (McDonald, Murphy, & Messinger, 2017; Newbigin et al., 2016). Concerning empathic responsiveness and prosocial behaviour in response to the experimenter's distress, no difference was found between autistic and typical participants (Butean et al., 2014; McDonald et al., 2017; McDonald & Messinger, 2012; Newbigin et al., 2016; Rieffe et al., 2021; Scheeren, Koot, Mundy, Mous, & Begeer, 2013; Zadok, Gordon, Navon, Rabin, & Golan, 2021), except in one study (Li et al., 2023). Li et al. (2023) found that autistic children displayed less prosocial behaviours than typical children, but also noted that there was a greater augmentation in prosocial behaviour over time in that group. Finally, several studies suggested that autistic participants looked as much at the experimenter as typical participants (Campbell et al., 2017; Charman et al., 1997; Rieffe et al., 2021; Tsang, Gillespie-Lynch, & Hutman, 2016), although two studies found that autistic participants looked less at the experimenter than typical participants (Hutman et al., 2010; Li et al., 2023).

Another variation of the distress task consisted of a scenario in which the experimenter's drawing got torn to pieces by another experimenter once they had finished drawing it. The two studies that used this task had contradictory results. One study suggested that autistic children showed less concern and looked less at the experimenter (Hobson, Harris, Garcia-Perez, & Hobson, 2009), whereas the other found that they showed as much concern and looked as often at the experimenter as typical children (Newbigin et al., 2016).

Finally, two studies asked the children's parents to describe their child's behaviour when their child was in the presence of another person displaying emotional reactions. These studies showed that the parents of autistic children described their child as being less concerned and exhibiting less empathic responsiveness and prosocial behaviours than the parents of typical children (Hudry & Slaughter, 2009; Scheeren et al., 2013).

3.7.2. The empathic accuracy task

Four studies used the empathic accuracy task, an ecological empathy task, to measure cognitive and affective empathy. This task measures cognitive empathy as the extent to which the participants' rating of the target's emotional state matches the target's own rating of their emotional state. There were no differences between the autistic and typical group in studies using the empathic accuracy task when participants were asked to continuously rate the target's emotional state while watching the video (McKenzie, Russell, Golm, & Fairchild, 2022; Oliver et al., 2024; Santiesteban et al., 2021; Vilas, Reniers, & Ludlow, 2021). When the participants were asked to rate the target's emotional state after watching the video, one study found that the autistic group was less accurate than the typical group (Santiesteban et al., 2021), whereas another study did not find any difference between the two groups (McKenzie et al., 2022).

Affective empathy is measured by the empathic accuracy task as the extent to which the participants' rating of their own emotional state matches the target's rating of their own emotional state. When participants were asked to continuously rate their emotional state while watching the video, one study did not find any difference between autistic and typical participants (Santiesteban et al., 2021). When the participants were asked to rate their emotional state after watching the video, one study found that autistic participants had less affective empathy than typical participants (Santiesteban et al., 2021), whereas another did not find any difference between the two groups (McKenzie et al., 2022).

Finally, concerning affect sharing (i.e., the extent to which the participants' rating of their own emotional state matches the participants' rating of the target's emotional state), no difference was found between autistic and typical participants, neither when the rating took place while watching the video or after watching the video (McKenzie et al., 2022; Santiesteban et al., 2021).

3.7.3. Self-other empathy task

When exposed to pictures of emotional faces or emotional videos and asked to indicate how they felt, autistic participants less often reported that they felt the same emotion as the person in the picture/video than typical participants (Greimel et al., 2010; Schneider et al., 2013; Schulte-Rüther et al., 2011), except in a study by Klapwijk et al. (2016), in which there were no group differences. By contrast, when asked to indicate the depicted person's emotional state, all but one study (Greimel et al., 2010) did not find any difference between the autistic and typical groups (Klapwijk et al., 2016; Schneider et al., 2013; Schulte-Rüther et al., 2011). More specifically, Greimel et al. (2010) showed that although there were no group differences when the portrayed emotion was strong, autistic individuals were less accurate than typical individuals when the emotional expression was weak.

One study asked participants to rate how pleasant a touch stimulation was for themselves or for the other participant (Hoffmann, Koehne, Steinbeis, Dziobek, & Singer, 2016). In the first condition, only one of the participants received a stimulation (individual condition). In the second condition, the simultaneous condition, both participants received either the same type of stimulation (congruent trial) or different types of stimulation (incongruent trial; e.g., one received a pleasant stimulation and the other, an unpleasant stimulation). They found that in the individual condition, autistic participants rated both the pleasant and unpleasant stimulations as less intense for themselves and the other person than did typical participants. However, in the simultaneous condition, there was no difference between the two groups when the differences between the incongruent and congruent trials for the other- and self-ratings were compared, which the authors interpreted as indicating intact self-other distinction in autism (Hoffmann et al., 2016).

3.7.4. Empathy for physical pain

Empathy for physical pain tasks can be separated into two categories: self-other empathy for pain tasks and the rating or detection of another person's pain. There were no differences between autistic and typical participants in the self-other empathy for pain tasks in which participants were asked to rate the unpleasantness and intensity of their own pain either in response to a painful stimulation or to a picture depicting a painful situation (Bird et al., 2010; De Coster et al., 2018; Jeon et al., 2024; Minio-Paluello, Baron-Cohen, Avenanti, Walsh, & Aglioti, 2009; Thaler et al., 2018). Similarly, there were no differences between the groups when they were asked to rate the unpleasantness and intensity of the other person's pain (Bird et al., 2010; De Coster et al., 2018; Jeon et al., 2024; Minio-Paluello et al., 2009; Thaler et al., 2018).

Concerning the studies measuring the detection of another person's pain, among those that used a rating scale, three studies found no differences between the autistic and typical participants pain ratings (Krach et al., 2015; Li, Decety, Hua, Li, & Yi, 2024; Stroth et al., 2019), whereas one study suggested that autistic participants rated the depicted situations as less painful than typical participants (Chen et al., 2017). On the contrary, two studies suggested that autistic participants were less accurate than typical participants when participants had to decide whether the picture depicted a painful situation or not (Gu et al., 2015; Lassalle et al., 2019), whereas one study did not find any differences between the groups (Fan et al., 2014).

4. Discussion

This systematic review and meta-analysis investigated cognitive, affective, and unidimensional empathy in autism and examined the effect of several moderators on the difference in empathy scores between autistic and typical participants. It then probed whether the use of different empathy measures had an impact on the results of the studies. Finally, the results of studies using more ecological empathy measures or neuroimaging paradigms were synthesized.

4.1. Empathy and autism

The results of this meta-analysis suggest that autistic individuals on average display less cognitive empathy ($g = -0.85$) and, to a notably lesser extent, affective empathy ($g = -0.17$) than typical individuals. Thus, these findings provide support for the empathic imbalance hypothesis of autism (Smith, 2009), which postulates that autism is characterized by a deficit in cognitive empathy but by normal or even heightened sensitivity to affective empathy. This hypothesis further proposes that an imbalance between cognitive and affective empathy could lead autistic individuals to be overwhelmed by the sharing of another's emotions, thus leading to personal distress rather than empathic concern (Smith, 2009). Results from the IRI questionnaire are consistent with this suggestion, as a dissociation between personal distress and empathic concern was found. Indeed, on average, autistic individuals reported less empathic concern ($g = -0.59$) but heightened personal distress ($g = 0.67$) relative to typical individuals. Therefore, autistic individuals could seem unempathetic, whereas they might instead be struggling to regulate their empathic arousal (Smith, 2009). These various subcomponents of affective empathy are thus important to consider when assessing empathy in an autistic sample.

4.2. What is the impact of different measures on assessing empathy in autism?

The present meta-analysis found notable differences in the effect sizes obtained from different empathy measures, as well as between different subscales of a single measure. For example, the EQ, the most frequently used unidimensional empathy measure, systematically showed lower reported empathy for autistic than typical individuals. Importantly, previous factor analyses found that the EQ had either one global empathy factor or three factors, comprising cognitive empathy, emotional reactivity, and social skills (Allison et al., 2011; Lawrence et al., 2004). Although social skills are evidently related to empathy, they are not conceptually a part of empathy per se. Furthermore, several authors consider that the EQ primarily measures cognitive empathy (e.g., Greenberg, Warrier, Allison, & Baron-Cohen, 2018). Thus, these two factors may explain the very high effect size obtained for the EQ and unidimensional empathy analysis, as this analysis mainly included studies using the EQ or one of its various forms.

On the other hand, results from the IRI, the most frequently used multidimensional empathy measure, suggested that autistic individuals had, on average, lower perspective taking, fantasy, and empathic concern, but higher personal distress. The lower score on the perspective taking subscale ($g = -0.99$), the subscale considered to be the most representative of cognitive empathy, is consistent with the past literature on cognitive empathy and theory of mind in autism (Baron-Cohen, 2000; Smith, 2009; although see Gernsbacher & Yergeau, 2019 for an alternative view). Fantasy appears to measure both cognitive and affective empathy (Davis, 1980, 1983; De Corte et al., 2007) and, accordingly, the obtained effect size ($g = -0.42$) was smaller than for perspective taking. Finally, concerning the two affective empathy subscales, findings suggest that, on average, typical individuals feel more concern for others in distressing situations than autistic individuals, but that autistic individuals are more distressed than typical individuals in a tense interpersonal situation. This dissociation may be explained by the difference in focus between these two subscales, as empathic concern assesses other-oriented feelings of concern and compassion, whereas personal distress concerns self-oriented emotional responses to another's negative experience (Davis, 1980). In addition, the empathic concern subscale may involve more cognitive processes, as several of its items require putting oneself in another's position (Grynberg et al., 2010). A multidimensional assessment of empathy is thus warranted in studies assessing empathy in autism.

4.3. Sex and age-related differences in empathy

Concerning the influence of sex on empathy, although it was not a significant moderator for cognitive empathy, the difference in unidimensional empathy between autistic and typical women was greater than that between autistic and typical men. One possible explanation is provided by the Extreme Male Brain theory (Baron-Cohen, 2002, 2009). This theory posits that typical females tend to have greater empathy than systematizing and that typical males tend to have greater systematizing than empathy, whereas autistic individuals, irrespective of sex, tend to have higher systematizing and lower empathy (Baron-Cohen, 2002, 2009). As such, it suggests that autistic females have a profile more akin to that of typical males than females (Baron-Cohen, 2009). However, this theory is heavily based on results obtained from the EQ, which may explain why this sex-related difference was only found for unidimensional empathy. Of note, no studies have assessed sex-related differences between autistic and typical individuals concerning affective empathy, which is a clear gap in the current literature. Moreover, several authors suggest that sex may influence the behavioural presentation of autism and developmental trajectory of autistic characteristics (e.g., Hull, Mandy, & Petrides, 2017; Lai & Szatmari, 2020). Thus, further studies using more diverse methodologies are warranted before drawing definitive conclusions on whether normative sex differences in empathy are present in autism.

In terms of age, the differences in empathy scores between autistic and typical individuals were comparable in children and adults for cognitive, affective, and unidimensional empathy. The few longitudinal studies conducted to date, most of which have focused on infants or preschoolers, generally find that although autistic children show lower global empathic responses to distress compared to low-risk or typical children, this group difference tends to remain consistent over time (Campbell et al., 2017; Hutman et al., 2010; Li et al., 2023; McDonald & Messinger, 2012; Paz et al., 2024). Notably, the only longitudinal study of school-aged children reported greater gains in cognitive empathy in autistic children compared to typical children, although their scores remained lower overall (Kouklari, Tsermentseli, & Monks, 2019), suggesting potentially different developmental trajectories in autism, particularly in later childhood and adolescence. Thus, longitudinal studies across different developmental stages are needed to clarify age-related differences between autistic and typical males and females.

4.4. The influence of alexithymia and callous-unemotional traits

Alexithymia did not moderate either cognitive or affective empathy. The alexithymia hypothesis of autism proposes that the emotional difficulties associated with autism may be mainly attributable to co-occurring alexithymia rather than autism per se (Bird & Cook, 2013). This hypothesis was not supported by the current meta-analysis. However, the few studies that directly matched autistic and typical individuals for degrees of alexithymia did not find differences in empathy between the two populations (Bernhardt et al., 2014; Bird et al., 2010; Oakley et al., 2016). This may imply that methodological choices, such as controlling for alexithymia in subsequent statistical analyses versus matching participants on alexithymia from the start, may have an impact on whether alexithymia is found to moderate empathy or not. Furthermore, a number of studies assessed the correlation between empathy and alexithymia measures independently of group membership and generally found moderate to strong negative correlations (Bird et al., 2010; Lombardo, Barnes, Wheelwright, & Baron-Cohen, 2007; Mul, Stagg, Herbelin, & Aspell, 2018; Santiesteban et al., 2021). Thus, matching participants for alexithymia and reporting the correlation between empathy and alexithymia measures is recommended in future studies to better differentiate the distinct effects of autism and alexithymia on empathy.

A limitation of this meta-analysis is that the influence of callous-unemotional traits could not be investigated. The few studies that directly compared autistic youth, youth with conduct disorder and elevated callous-unemotional traits, and typical youth suggest the presence of different empathy profiles in these conditions (Klapwijk et al., 2016; Schwenck et al., 2012). Moreover, a study found that when controlling for callous-unemotional traits, there no longer were differences in affective empathy between the autistic and typical group, whereas the cognitive empathy deficits remained (Tkalcec et al., 2023). As there is a higher occurrence of callous-unemotional traits in the autistic population than in the general population (Maguire, Warman, Blumenfeld, & Langdon, 2024), further studies could examine the relationship between autism and empathy while considering the impact of these traits.

4.5. Methodological considerations

Another major limitation of this meta-analysis is that it was limited by the definitional issues surrounding empathy. There are numerous definitions for empathy, and the existing empathy measures were all based on different definitions. Thus, these measures may not be directly comparable. Furthermore, there sometimes is a mismatch between the empathy definition offered in the studies' introduction and the way empathy is operationalized in the chosen measure (Bollen, 2023). Similarly, although empathy conceptualizations evolved over time, this is not reflected in the empathy measures chosen to assess empathy in autistic individuals. Indeed, several newer empathy measures, usually more aligned with current empathy conceptualizations and assessing several cognitive and affective empathy components, were either underrepresented or not employed in the articles that were included in the current meta-analysis. For instance, only three articles used the Multifaceted Empathy Test (MET; Dziobek et al., 2008), and six articles used the Questionnaire of Affective and Cognitive Empathy (QCAE; Reniers, Corcoran, Drake, Shryane, & Vollm, 2011). Similarly, none of the articles used empathy measures such as the Affective and Cognitive Measure of Empathy (ACME; Vachon & Lynam, 2016), the Empathy Components Questionnaire (ECQ; Batchelder, Brosnan, & Ashwin, 2017), or the Measure of Empathy in Early Childhood (MEEC; Kimonis, Jain, Neo, Fleming, & Briggs, 2023).

In addition, most studies relied on self-report measures, which reflect people's perception of their empathic abilities rather than their true abilities. For example, participant scores on the perspective taking subscale of the IRI are only weakly associated with their performance on the RMET ($rs = 0.11$ to 0.12) (Murphy & Lilienfeld, 2019; Sunahara et al., 2022), although both measures purportedly assess cognitive empathy. Self-report questionnaires are also prone to several biases, such as social desirability (i.e., presenting oneself in an overly positive light) and extreme responding (i.e., selecting answers at the extremities of a rating scale), some of which may have a differential impact on typical and autistic groups (Paulhus & Vazire, 2007; Sher & Oliver, 2022). Moreover, certain items might be ambiguous or use non-literal language, which can affect the way autistic individuals interpret and answer these items (Harrison, Brownlow, Ireland, & Piovesana, 2022; Sher & Oliver, 2022). A COSMIN review on empathy self-report questionnaires found that content and criterion validity, as well as measurement invariance and error, were never assessed for the EQ or IRI on autistic individuals (Harrison et al., 2022). Furthermore, these questionnaires had a potentially high risk of bias, and their psychometric properties were generally deemed insufficient (Harrison et al., 2022). Thus, both the EQ and its various forms and the IRI were deemed to require further research before being used either in research or clinical settings with autistic individuals. Moreover, the same review deemed that the IRI required further research to validate its use for typical individuals, whereas the EQ was simply “not recommended for use”. This is particularly problematic when considering that most studies that have assessed empathy in autism used these questionnaires. Of note, to date, only one study assessed measurement invariance for an empathy measure in autism (Brett, Preece, Becerra, Whitehouse, & Maybery, 2024; Harrison et al., 2022). Brett et al. (2024) indeed showed that partial invariance was supported for the Perth Empathy Scale (PES), which suggests that it assesses empathy similarly in autistic and non-autistic populations and may validly be used to compare both groups.

Finally, the empathy difficulties noted in autism were observed using empathy measures generally developed by and for typical individuals. However, one of the ramifications of the double empathy problem is that neurotypical individuals might find it as hard to understand the emotional states of autistic individuals as autistic individuals find it hard to understand those of neurotypical individuals (Chown, 2014; Milton, 2012). In this regard, several studies showed that typical individuals indeed have more difficulties understanding the mental and emotional states of autistic individuals than of typical individuals (e.g., Edey et al., 2016; Jacques, Courchesne, Mineau, Dawson, & Mottron, 2022; Sheppard, Pillai, Wong, Ropar, & Mitchell, 2016). Moreover, among neurotypical individuals, both cognitive and affective empathy are higher for in-group members (Eres & Molenberghs, 2013). To date, only two studies by the same research group seem to have assessed this effect in the context of autism. These studies show that autistic individuals indeed empathize more and have greater brain activation in several regions associated with empathy than typical individuals when reading sentences describing autistic characters whereas typical individuals empathize more and have greater brain activation in empathy-related regions when reading sentences describing non-autistic characters (Komeda et al., 2015; Komeda et al., 2019). Thus, these results raise the possibility that a mismatch in neurotype between two individuals may negatively impact empathy.

To sum up, finer points regarding the different cognitive and affective empathy components and the impact of situational factors, such as group membership, have been understudied in autism. As such, the use of both self-report and behavioural measures in line with current empathy conceptualizations and of more naturalistic tasks assessing empathy in autistic, typical, and mixed dyads is recommended to gain a more nuanced picture.

4.6. Statistical considerations

From a statistical point of view, most studies had a small sample size and several analyses in the current meta-analysis showed very high heterogeneity. This is a pervasive issue in psychological research, as shown by Stanley, Carter, and Doucouliagos (2018), which found that only around 8 % of the studies included in the 200 meta-analyses that they assessed were adequately powered and that the median heterogeneity (I^2) was high. Furthermore, Cristea, Georgescu, and Ioannidis (2021) determined that, on average, there was a greater difference in the effect size between the most cited study and the study with the largest sample size than between the most cited study and the corresponding meta-analysis within several emotion-related domains. Similarly, Kvarven, Stromland, and Johannesson (2020) found that the effect sizes obtained from meta-analyses were, on average, almost three times higher than those of matching multisite preregistered replication studies. As such, the summary effects obtained in the current meta-analysis may have been inflated by several factors and should be interpreted with caution.

The findings that autistic individuals on average have both lower cognitive and affective empathy relative to typical individuals could lead to the conclusion that they have a global empathy deficit. However, greater relative variability in empathy scores in autistic individuals was found, which suggests that some autistic individuals may have empathy scores that are similar or even superior to those of typical individuals, whereas others may have more pronounced difficulties. Moreover, a recent study showed that, for data with a normal distribution, even when a large effect size of $d = 1.00$ was obtained, 48 % of autistic individuals were, on average, still within one standard deviation of the mean of the typical participants, whereas this proportion increased to 67 % of autistic individuals for a small effect size ($d = 0.20$) (Loth et al., 2021). Indeed, for a measure to have 80 % sensitivity and specificity, the required benchmark for diagnostic utility, an effect size of 1.66 must be reached. As such, from a statistical point of view, even in the presence of a large effect size, almost half of autistic individuals do not have a deficit on the measured construct (Loth et al., 2021). This, together with the greater variability in empathy scores observed in autistic individuals, suggests that empathy might be of limited clinical relevance for an autism diagnosis and further emphasizes the need for caution when interpreting group-level empathy differences in autism.

5. Conclusion

A comprehensive overview of the literature on empathy in autism suggests that, on average, autistic individuals have a different empathy profile than typical individuals, characterized by reduced cognitive empathy and limited difficulties with the affective empathy component. By contrast, studies using unidimensional empathy measures have almost systematically reported autistic individuals to show reduced empathy. Furthermore, the results highly rely on the measure chosen to assess empathy in autism, as there were noticeable differences in the effect sizes obtained depending on the chosen measure. Finally, several methodological and statistical issues in studies investigating empathy in autism were noted. Thus, the taken for granted conclusions on the autistic deficit in empathy must be treated with caution as long as research using measures more aligned with current empathy conceptualizations is not conducted.

Funding

This work was supported by a Canadian Institutes of Health Research (CIHR) masters award to Noémie Cusson, by a Fonds de Recherche du Québec – Santé (FRQS) career award to Isabelle Soulières, and by the Chaire de Recherche Marcel et Rolande Gosselin en neurosciences cognitives de l’autisme de l’Université de Montréal held by Laurent Mottron.

Declaration of competing interest

The authors declare the following financial interests/personal relationships which may be considered as potential competing interests:

Corresponding author receives royalties from books and some teaching. If there are other authors, they declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgements

We thank Amélie Girard for providing research assistance and Alex Edelman & Associates for English editing.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.cpr.2025.102623.

References

Allison, C., Baron-Cohen, S., Wheelwright, S. J., Stone, M. H., & Muncer, S. J. (2011). Psychometric analysis of the Empathy Quotient (EQ). *Personality and Individual Differences, 51*, 829–835. https://doi.org/10.1016/j.paid.2011.07.005

American Psychiatric Association. (2022). *Diagnostic and statistical manual of mental disorders* (5th ed., text rev.). https://doi.org/10.1176/appi.books.9780890425787

Audras-Torrent, L., Miniarikova, E., Couty, F., Dellapiazza, F., Berard, M., Michelon, C., … Baghdadli, A. (2021). WISC-V profiles and their correlates in children with autism spectrum disorder without intellectual developmental disorder: Report from the ELENA cohort. *Autism Research, 14*, 997–1006. https://doi.org/10.1002/aur.2444

Avenanti, A., Minio-Paluello, I., Bufalari, I., & Aglioti, S. M. (2006). Stimulus-driven modulation of motor-evoked potentials during observation of others' pain. *Neuroimage, 32*, 316–324. https://doi.org/10.1016/j.neuroimage.2006.03.010

Banzhaf, C., Hoffmann, F., Kanske, P., Fan, Y., Walter, H., Spengler, S., … Bermpohl, F. (2018). Interacting and dissociable effects of alexithymia and depression on empathy. *Psychiatry Research, 270*, 631–638. https://doi.org/10.1016/j.psychres.2018.10.045

Baron-Cohen, S. (2000). Theory of mind and autism: A fifteen year review. In S. Baron-Cohen, H. Tager-Flusberg, & D. J. Cohen (Eds.), *Understanding other minds: Perspectives from developmental cognitive neuroscience* (2 ed., pp. 3–20). New York, NY, US: Oxford University Press.

Baron-Cohen, S. (2002). The extreme male brain theory of autism. *Trends in Cognitive Sciences, 6*, 248–254. https://doi.org/10.1016/S1364-6613(02)01904-6

Baron-Cohen, S. (2009). Autism: The empathizing-systemizing (E-S) theory. *Annals of the New York Academy of Sciences, 1156*, 68–80. https://doi.org/10.1111/j.1749-6632.2009.04467.x

Baron-Cohen, S. (2010). Empathizing, systemizing, and the extreme male brain theory of autism. *Progress in Brain Research, 186*, 167–175. https://doi.org/10.1016/b978-0-444-53630-3.00011-7

Baron-Cohen, S., Knickmeyer, R. C., & Belmonte, M. K. (2005). Sex differences in the brain: Implications for explaining autism. *Science, 310*, 819–823. https://doi.org/10.1126/science.1115455

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*, 241–251. https://doi.org/10.1111/1469-7610.00715

Baron-Cohen, S., Wheelwright, S., Robinson, J., & Woodbury-Smith, M. (2005). The Adult Asperger Assessment (AAA): A diagnostic method. *Journal of Autism and Developmental Disorders, 35*, 807–819. https://doi.org/10.1007/s10803-005-0026-5

Baron-Cohen, S., & Wheelwright, S. J. (2004). The Empathy Quotient: An investigation of adults with Asperger syndrome or high functioning autism, and normal sex differences. *Journal of Autism and Developmental Disorders, 34*, 163–175. https://doi.org/10.1023/B:JADD.0000022607.19833.00

Batchelder, L., Brosnan, M., & Ashwin, C. (2017). The development and validation of the Empathy Components Questionnaire (ECQ). *PLoS One, 12*, Article e0169185. https://doi.org/10.1371/journal.pone.0169185

Bermond, B., Clayton, K., Liberova, A., Luminet, O., Maruszewski, T., Ricci Bitti, P. E., … Wicherts, J. (2007). A cognitive and an affective dimension of alexithymia in six languages and seven populations. *Cognition & Emotion, 21*, 1125–1136. https://doi.org/10.1080/02699930601056989

Bernhardt, B. C., & Singer, T. (2012). The neural basis of empathy. *Annual Review of Neuroscience, 35*, 1–23. https://doi.org/10.1146/annurev-neuro-062111-150536

Bernhardt, B. C., Valk, S. L., Silani, G., Bird, G., Frith, U., & Singer, T. (2014). Selective disruption of sociocognitive structural brain networks in autism and alexithymia. *Cerebral Cortex, 24*, 3258–3267. https://doi.org/10.1093/cercor/bht182

Bird, G., & Cook, R. (2013). Mixed emotions: The contribution of alexithymia to the emotional symptoms of autism. *Translational Psychiatry, 3*, Article e285. https://doi.org/10.1038/tp.2013.61

Bird, G., Silani, G., Brindley, R., White, S., Frith, U., & Singer, T. (2010). Empathic brain responses in insula are modulated by levels of alexithymia but not autism. *Brain, 133*, 1515–1525. https://doi.org/10.1093/brain/awq060

Bird, G., & Viding, E. (2014). The self to other model of empathy: Providing a new framework for understanding empathy impairments in psychopathy, autism, and alexithymia. *Neuroscience and Biobehavioral Reviews, 47*, 520–532. https://doi.org/10.1016/j.neubiorev.2014.09.021

Bollen, C. (2023). A reflective guide on the meaning of empathy in autism research. *Methods in Psychology, 8*. https://doi.org/10.1016/j.metip.2022.100109

Bons, D., van den Broek, E., Scheepers, F., Herpers, P., Rommelse, N., & Buitelaar, J. K. (2013). Motor, emotional, and cognitive empathy in children and adolescents with autism spectrum disorder and conduct disorder. *Journal of Abnormal Child Psychology, 41*, 425–443. https://doi.org/10.1007/s10802-012-9689-5

Borenstein, M., Hedges, L. V., Higgins, J. P., & Rothstein, H. R. (2010). A basic introduction to fixed-effect and random-effects models for meta-analysis. *Research Synthesis Methods, 1*, 97–111. https://doi.org/10.1002/jrsm.12

Brett, J. D., Preece, D. A., Becerra, R., Whitehouse, A., & Maybery, M. T. (2024). Empathy and autism: Establishing the structure and different manifestations of empathy in autistic individuals using the Perth Empathy Scale. *Journal of Autism and Developmental Disorders*. https://doi.org/10.1007/s10803-024-06491-3

Butean, I., Costescu, C., & Dobrean, A. (2014). Differences between empathic responses in children with autism spectrum disorder and typically developing children. *Journal of Evidence-Based Psychotherapies, 14*, 197–209.

Campbell, S. B., Moore, E. L., Northrup, J., & Brownell, C. A. (2017). Developmental changes in empathic concern and self-understanding in toddlers at genetic risk for autism spectrum disorder. *Journal of Autism and Developmental Disorders, 47*, 2690–2702. https://doi.org/10.1007/s10803-017-3192-3

Charman, T., Swettenham, J., Baron-Cohen, S., Cox, A., Baird, G., & Drew, A. (1997). Infants with autism: An investigation of empathy, pretend play, joint attention, and imitation. *Developmental Psychology, 33*, 781–789. https://doi.org/10.1037//0012-1649.33.5.781

Chen, C., Hung, A. Y., Fan, Y. T., Tan, S., Hong, H., & Cheng, Y. (2017). Linkage between pain sensitivity and empathic response in adolescents with autism spectrum conditions and conduct disorder symptoms. *Autism Research, 10*, 267–275. https://doi.org/10.1002/aur.1653

Cheung, M. W. (2019). A guide to conducting a meta-analysis with non-independent effect sizes. *Neuropsychology Review, 29*, 387–396. https://doi.org/10.1007/s11065-019-09415-6

Chown, N. (2014). More on the ontological status of autism and double empathy. *Disability and Society, 29*, 1672–1676. https://doi.org/10.1080/09687599.2014.949625

Christov-Moore, L., Simpson, E. A., Coude, G., Grigaityte, K., Iacoboni, M., & Ferrari, P. F. (2014). Empathy: Gender effects in brain and behavior. *Neuroscience and Biobehavioral Reviews, 46*, 604–627. https://doi.org/10.1016/j.neubiorev.2014.09.001

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2 ed.). Hillsdale, NJ: Lawrence Erlbaum Associates.

Cristea, I. A., Georgescu, R., & Ioannidis, J. P. A. (2021). Effect sizes reported in highly cited emotion research compared with larger studies and meta-analyses addressing the same questions. *Clinical Psychological Science*. https://doi.org/10.1177/21677026211049366

Cuff, B. M. P., Brown, S. J., Taylor, L., & Howat, D. J. (2016). Empathy: A review of the concept. *Emotion Review, 8*, 144–153. https://doi.org/10.1177/1754073914558466

Davis, M. H. (1980). A multidimensional approach to individual differences in empathy. *JSAS Catalog of Selected Documents in Psychology*, 10.

Davis, M. H. (1983). Measuring individual differences in empathy: Evidence for a multidimensional approach. *Journal of Personality and Social Psychology, 44*, 113–126.

De Corte, K., Buysse, A., Verhofstadt, L., Roeyers, H., Ponnet, K., & Davis, M. H. (2007). Measuring empathic tendencies: Reliability and validity of the Dutch version of the Interpersonal Reactivity Index. *Psychologica Belgica, 47*, 235–260. https://doi.org/10.5334/pb-47-4-235

De Coster, L., Wiersema, J. R., Deschrijver, E., & Brass, M. (2018). The effect of being imitated on empathy for pain in adults with high-functioning autism: Disturbed self-other distinction leads to altered empathic responding. *Autism, 22*, 712–727. https://doi.org/10.1177/1362361317701268

Deeks, J. J., Higgins, J. P. T., & Altman, D. G. (2019). Chapter 10: Analysing data and undertaking meta-analyses. In J. P. T. Higgins, J. Thomas, J. Chandler, M. Cumpston, T. Li, M. J. Page, & V. A. Welch (Eds.), *Vol. 6.0. Cochrane handbook for systematic reviews of interventions*. Cochrane.

Duval, S., & Tweedie, R. (2000a). A Nonparametric "Trim and Fill" method of accounting for publication bias in meta-analysis. Journal of the American Statistical Association, 95, 89–98, https://doi.org/10.1080/01621459.2000.10473905.

Duval, S., & Tweedie, R. (2000b). Trim and fill: A simple funnel-plot-based method of testing and adjusting for publication bias in meta-analysis. *Biometrics, 56*, 455–463. https://doi.org/10.1111/j.0006-341X.2000.00455.x

Dvash, J., & Shamay-Tsoory, S. G. (2014). Theory of mind and empathy as multidimensional constructs. *Topics in Language Disorders, 34*, 282–295. https://doi.org/10.1097/tld.0000000000000040

Dziobek, I., Rogers, K., Fleck, S., Bahnemann, M., Heekeren, H. R., Wolf, O. T., & Convit, A. (2008). Dissociation of cognitive and emotional empathy in adults with Asperger syndrome using the Multifaceted Empathy Test (MET). *Journal of Autism and Developmental Disorders, 38*, 464–473. https://doi.org/10.1007/s10803-007-0486-x

Edey, R., Cook, J., Brewer, R., Johnson, M. H., Bird, G., & Press, C. (2016). Interaction takes two: Typical adults exhibit mind-blindness towards those with autism spectrum disorder. *Journal of Abnormal Psychology, 125*, 879–885. https://doi.org/10.1037/abn0000199

Egger, M., Smith, G. D., Schneider, M., & Minder, C. (1997). Bias in meta-analysis detected by a simple, graphical test. *British Medical Journal, 315*, 629–634. https://doi.org/10.1136/bmj.315.7109.629

Engelen, E.-M., & Röttger-Rössler, B. (2012). Current disciplinary and interdisciplinary debates on empathy. *Emotion Review, 4*, 3–8. https://doi.org/10.1177/1754073911422287

Eres, R., & Molenberghs, P. (2013). The influence of group membership on the neural correlates involved in empathy. *Frontiers in Human Neuroscience, 7*. https://doi.org/10.3389/fnhum.2013.00176

Fan, Y., Chen, C., Chen, S. C., Decety, J., & Cheng, Y. (2014). Empathic arousal and social understanding in individuals with autism: Evidence from fMRI and ERP measurements. *Social Cognitive and Affective Neuroscience, 9*, 1203–1213. https://doi.org/10.1093/scan/nst101

Fan, Y., Duncan, N. W., de Greck, M., & Northoff, G. (2011). Is there a core neural network in empathy? An fMRI based quantitative meta-analysis. *Neuroscience and Biobehavioral Reviews, 35*, 903–911. https://doi.org/10.1016/j.neubiorev.2010.10.009

Fatima, M., & Babu, N. (2023). Cognitive and affective empathy in autism spectrum disorders: A meta-analysis. *Review Journal of Autism and Developmental Disorders*. https://doi.org/10.1007/s40489-023-00364-8

Fernández-Castilla, B., Declercq, L., Jamshidi, L., Beretvas, S. N., Onghena, P., & Van den Noortgate, W. (2020). Visual representations of meta-analyses of multiple outcomes: Extensions to forest plots, funnel plots, and caterpillar plots. *Methodology, 16*, 299–315. https://doi.org/10.5964/meth.4013

Gernsbacher, M. A., & Yergeau, M. (2019). Empirical failures of the claim that autistic people lack a theory of mind. *Archives of Scientific Psychology, 7*, 102–118. https://doi.org/10.1037/arc0000067

Gillberg, C. L. (1992). Autism and autistic-like conditions: Subclasses among disorders of empathy. *Journal of Child Psychology and Psychiatry, 33*, 813–842. https://doi.org/10.1111/j.1469-7610.1992.tb01959.x

Greenberg, D. M., Warrier, V., Allison, C., & Baron-Cohen, S. (2018). Testing the empathizing-systemizing theory of sex differences and the extreme male brain theory of autism in half a million people. *Proceedings of the National Academy of Sciences of the United States of America, 115*, 12152–12157. https://doi.org/10.1073/pnas.1811032115

Greimel, E., Schulte-Ruther, M., Kircher, T., Kamp-Becker, I., Remschmidt, H., Fink, G. R., … Konrad, K. (2010). Neural mechanisms of empathy in adolescents with autism spectrum disorder and their fathers. *Neuroimage, 49*, 1055–1065. https://doi.org/10.1016/j.neuroimage.2009.07.057

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*, 1–10. https://doi.org/10.1186/2040-2392-5-42

Grynberg, D., Luminet, O., Corneille, O., Grèzes, J., & Berthoz, S. (2010). Alexithymia in the interpersonal domain: A general deficit of empathy? *Personality and Individual Differences, 49*, 845–850. https://doi.org/10.1016/j.paid.2010.07.013

Gu, X., Eilam-Stock, T., Zhou, T., Anagnostou, E., Kolevzon, A., Soorya, L., … Fan, J. (2015). Autonomic and brain responses associated with empathy deficits in autism spectrum disorder. *Human Brain Mapping, 36*, 3323–3338. https://doi.org/10.1002/hbm.22840

Harmsen, I. E. (2019). Empathy in autism spectrum disorder. *Journal of Autism and Developmental Disorders, 49*, 3939–3955. https://doi.org/10.1007/s10803-019-04087-w

Harrer, M., Cuijpers, P., Furukawa, T., & Ebert, D. D. (2019). Dmetar: Companion R Package For The Guide 'Doing Meta-Analysis in R' (Version 0.0.9000). http://dmetar.protectlab.org/.

Harrison, J. L., Brownlow, C. L., Ireland, M. J., & Piovesana, A. M. (2022). Empathy measurement in autistic and nonautistic adults: A COSMIN systematic literature review. *Assessment, 29*, 332–350. https://doi.org/10.1177/1073191120964564

Hartung, J., & Knapp, G. (2001). On tests of the overall treatment effect in meta-analysis with normally distributed responses. *Statistics in Medicine, 20*, 1771–1782. https://doi.org/10.1002/sim.791

Henry, J. D., von Hippel, W., Molenberghs, P., Lee, T., & Sachdev, P. S. (2016). Clinical assessment of social cognitive function in neurological disorders. *Nature Reviews. Neurology, 12*, 28–39. https://doi.org/10.1038/nrneurol.2015.229

Higgins, J. P. T., Thomas, J., Chandler, J., Cumpston, M., Li, T., Page, M. J., & Welch, V. A. (2022). Cochrane handbook for systematic reviews of interventions (version 6.3). www.training.cochrane.org/handbook.

Higgins, J. P. T., Thompson, S. G., Deeks, J. J., & Altman, D. G. (2003). Measuring inconsistency in meta-analyses. *British Medical Journal, 327*, 557–560. https://doi.org/10.1136/bmj.327.7414.557

Hobson, J. A., Harris, R., Garcia-Perez, R., & Hobson, R. P. (2009). Anticipatory concern: A study in autism. *Developmental Science, 12*, 249–263. https://doi.org/10.1111/j.1467-7687.2008.00762.x

Hoffman, M. L. (2000). Development of empathic distress. In *Empathy and moral development: Implications for caring and justice* (pp. 63–92). Cambrige, United Kingdom: Cambrige University Press.

Hoffmann, F., Koehne, S., Steinbeis, N., Dziobek, I., & Singer, T. (2016). Preserved self-other distinction during empathy in autism is linked to network integrity of right supramarginal gyrus. *Journal of Autism and Developmental Disorders, 46*, 637–648. https://doi.org/10.1007/s10803-015-2609-0

Hudry, K., & Slaughter, V. (2009). Agent familiarity and emotional context influence the everyday empathic responding of young children with autism. *Research in Autism Spectrum Disorders, 3*, 74–85. https://doi.org/10.1016/j.rasd.2008.04.004

Hull, L., Mandy, W., & Petrides, K. V. (2017). Behavioural and cognitive sex/gender differences in autism spectrum condition and typically developing males and females. *Autism, 21*, 706–727. https://doi.org/10.1177/1362361316669087

Hutman, T., Rozga, A., DeLaurentis, A. D., Barnwell, J. M., Sugar, C. A., & Sigman, M. (2010). Response to distress in infants at risk for autism: A prospective longitudinal study. *Journal of Child Psychology and Psychiatry, and Allied Disciplines, 51*, 1010–1020. https://doi.org/10.1111/j.1469-7610.2010.02270.x

Ickes, W., Stinson, L., Bissonnette, V., & Garcia, S. (1990). Naturalistic social cognition: Empathic accuracy in mixed-sex dyads. *Journal of Personality and Social Psychology, 59*, 730–742.

IntHout, J., Ioannidis, J. P., Rovers, M. M., & Goeman, J. J. (2016). Plea for routinely presenting prediction intervals in meta-analysis. *BMJ Open, 6*, Article e010247. https://doi.org/10.1136/bmjopen-2015-010247

Jacques, C., Courchesne, V., Mineau, S., Dawson, M., & Mottron, L. (2022). Positive, negative, neutral-or unknown? The perceived valence of emotions expressed by young autistic children in a novel context suited to autism. *Autism, 26*, 1833–1848. https://doi.org/10.1177/13623613211068221

Jeon, H., Hur, A., Lee, H., Shin, Y. W., Lee, S. I., Shin, C. J., … Son, J. W. (2024). The relationship between brain activation for taking others' perspective and interoceptive abilities in autism spectrum disorder: An fMRI study. *Journal of the Korean Academy of Child and Adolescent Psychiatry, 35*, 197–209. https://doi.org/10.5765/jkacap.240008

Kimonis, E. R., Jain, N., Neo, B., Fleming, G. E., & Briggs, N. (2023). Development of an empathy rating scale for young children. *Assessment, 30*, 37–50. https://doi.org/10.1177/10731911211038629

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

Kittel, A. F. D., Olderbak, S., & Wilhelm, O. (2022). Sty in the mind's eye: A meta-analytic investigation of the nomological network and internal consistency of the "Reading the Mind in the Eyes" test. *Assessment, 29*, 872–895. https://doi.org/10.1177/1073191121996469

Klapwijk, E. T., Aghajani, M., Colins, O. F., Marijnissen, G. M., Popma, A., van Lang, N. D., … Vermeiren, R. R. (2016). Different brain responses during empathy in autism spectrum disorders versus conduct disorder and callous-unemotional traits. *Journal of Child Psychology and Psychiatry, and Allied Disciplines, 57*, 737–747. https://doi.org/10.1111/jcpp.12498

Knapp, G., & Hartung, J. (2003). Improved tests for a random effects meta-regression with a single covariate. *Statistics in Medicine, 22*, 2693–2710. https://doi.org/10.1002/sim.1482

Kogler, L., Muller, V. I., Werminghausen, E., Eickhoff, S. B., & Derntl, B. (2020). Do I feel or do I know? Neuroimaging meta-analyses on the multiple facets of empathy. *Cortex, 129*, 341–355. https://doi.org/10.1016/j.cortex.2020.04.031

Kok, F. M., Groen, Y., Becke, M., Fuermaier, A. B., & Tucha, O. (2016). Self-reported empathy in adult women with autism spectrum disorders - A systematic mini review. *PLoS One, 11*, Article e0151568. https://doi.org/10.1371/journal.pone.0151568

Komeda, H., Kosaka, H., Fujioka, T., Jung, M., & Okazawa, H. (2019). Do individuals with autism spectrum disorders help other people with autism spectrum disorders? An investigation of empathy and helping motivation in adults with autism spectrum disorder. *Frontiers in Psychiatry, 10*, 376. https://doi.org/10.3389/fpsyt.2019.00376

Komeda, H., Kosaka, H., Saito, D. N., Mano, Y., Jung, M., Fujii, T., … Okazawa, H. (2015). Autistic empathy toward autistic others. *Social Cognitive and Affective Neuroscience, 10*, 145–152. https://doi.org/10.1093/scan/nsu126

Kouklari, E.-C., Tsermentseli, S., & Monks, C. P. (2019). Developmental trends of hot and cool executive function in school-aged children with and without autism spectrum disorder: Links with theory of mind. *Development and Psychopathology, 31*, 541–556. https://doi.org/10.1017/S0954579418000081

Krach, S., Kamp-Becker, I., Einhauser, W., Sommer, J., Frassle, S., Jansen, A., … Paulus, F. M. (2015). Evidence from pupillometry and fMRI indicates reduced neural response during vicarious social pain but not physical pain in autism. *Human Brain Mapping, 36*, 4730–4744. https://doi.org/10.1002/hbm.22949

Kvarven, A., Stromland, E., & Johannesson, M. (2020). Comparing meta-analyses and preregistered multiple-laboratory replication projects. *Nature Human Behavior, 4*, 423–434. https://doi.org/10.1038/s41562-019-0787-z

Lai, M. C., & Szatmari, P. (2020). Sex and gender impacts on the behavioural presentation and recognition of autism. *Current Opinion in Psychiatry, 33*, 117–123. https://doi.org/10.1097/YCO.0000000000000575

Lamm, C., Decety, J., & Singer, T. (2011). Meta-analytic evidence for common and distinct neural networks associated with directly experienced pain and empathy for pain. *Neuroimage, 54*, 2492–2502. https://doi.org/10.1016/j.neuroimage.2010.10.014

Langan, D., Higgins, J. P. T., Jackson, D., Bowden, J., Veroniki, A. A., Kontopantelis, E., … Simmonds, M. (2019). A comparison of heterogeneity variance estimators in simulated random-effects meta-analyses. *Research Synthesis Methods, 10*, 83–98. https://doi.org/10.1002/jrsm.1316

Lassalle, A., Zurcher, N. R., Porro, C. A., Benuzzi, F., Hippolyte, L., Lemonnier, E., … Hadjikhani, N. (2019). Influence of anxiety and alexithymia on brain activations associated with the perception of others' pain in autism. *Social Neuroscience, 14*, 359–377. https://doi.org/10.1080/17470919.2018.1468358

Lawrence, E. J., Shaw, P., Baker, D., Baron-Cohen, S., & David, A. S. (2004). Measuring empathy: Reliability and validity of the Empathy Quotient. *Psychological Medicine, 34*, 911–919. https://doi.org/10.1017/s0033291703001624

Li, B., Blijd-Hoogewys, E., Stockmann, L., Vergari, I., & Rieffe, C. (2023). Toward feeling, understanding, and caring: The development of empathy in young autistic children. *Autism, 27*, 1204–1218. https://doi.org/10.1177/13623613221117955

Li, T., Decety, J., Hua, Z., Li, G., & Yi, L. (2024). Empathy in autistic children: Emotional overarousal in response to others' physical pain. *Autism Research, 17*, 1640–1650. https://doi.org/10.1002/aur.3200

Lombardo, M. V., Barnes, J. L., Wheelwright, S. J., & Baron-Cohen, S. (2007). Self-referential cognition and empathy in autism. *PLoS One, 2*, Article e883. https://doi.org/10.1371/journal.pone.0000883

Lord, C., Rutter, M., DiLavore, P. C., Risi, S., Gotham, K., & Bishop, S. L. (2012). *Autism diagnostic observation schedule–2nd edition (ADOS-2)*. Torrance, CA: Western Psychological Services.

Loth, E., Ahmad, J., Chatham, C., Lopez, B., Carter, B., Crawley, D., … Dumas, G. (2021). The meaning of significant mean group differences for biomarker discovery. *PLoS Computational Biology, 17*, Article e1009477. https://doi.org/10.1371/journal.pcbi.1009477

Lüdecke, D. (2019). Esc: Effect Size Computation for Meta Analysis (Version 0.5.1). https://CRAN.R-project.org/package=esc.

Maguire, K., Warman, H., Blumenfeld, F., & Langdon, P. E. (2024). The relationship between psychopathy and autism: A systematic review and narrative synthesis. *Frontiers in Psychiatry, 15*. https://doi.org/10.3389/fpsyt.2024.1375170

Mathersul, D., McDonald, S., & Rushby, J. A. (2013). Understanding advanced theory of mind and empathy in high functioning adults with autism spectrum disorder. *Journal of Clinical and Experimental Neuropsychology, 35*, 655–668. https://doi.org/10.1080/13803395.2013.809700

McDonald, N. M., & Messinger, D. (2011). *The development of empathy: How, when, and why. Moral behavior and free will: A neurobiological and philosophical approach* (pp. 333–359).

McDonald, N. M., & Messinger, D. S. (2012). Empathic responding in toddlers at risk for an autism spectrum disorder. *Journal of Autism and Developmental Disorders, 42*, 1566–1573. https://doi.org/10.1007/s10803-011-1390-y

McDonald, N. M., Murphy, H. G., & Messinger, D. S. (2017). Empathic responding in preschool-aged children with familial risk for autism. *Autism Research, 10*, 1621–1628. https://doi.org/10.1002/aur.1819

McKenzie, K., Russell, A., Golm, D., & Fairchild, G. (2022). Empathic accuracy and cognitive and affective empathy in young adults with and without autism spectrum disorder. *Journal of Autism and Developmental Disorders, 52*, 2004–2018. https://doi.org/10.1007/s10803-021-05093-7

McShane, B. B., Bockenholt, U., & Hansen, K. T. (2016). Adjusting for publication bias in meta-analysis: An evaluation of selection methods and some cautionary notes. *Perspectives on Psychological Science, 11*, 730–749. https://doi.org/10.1177/1745691616662243

Milton, D. E. M. (2012). On the ontological status of autism: The 'double empathy problem'. *Disability & Society, 27*, 883–887. https://doi.org/10.1080/09687599.2012.710008

Minio-Paluello, I., Baron-Cohen, S., Avenanti, A., Walsh, V., & Aglioti, S. M. (2009). Absence of embodied empathy during pain observation in Asperger syndrome. *Biological Psychiatry, 65*, 55–62. https://doi.org/10.1016/j.biopsych.2008.08.006

Mottron, L., & Bzdok, D. (2020). Autism spectrum heterogeneity: Fact or artifact? *Molecular Psychiatry, 25*, 3178–3185. https://doi.org/10.1038/s41380-020-0748-y

Mul, C. L., Stagg, S. D., Herbelin, B., & Aspell, J. E. (2018). The feeling of me feeling for you: Interoception, alexithymia and empathy in autism. *Journal of Autism and Developmental Disorders, 48*, 2953–2967. https://doi.org/10.1007/s10803-018-3564-3

Murphy, B. A., & Lilienfeld, S. O. (2019). Are self-report cognitive empathy ratings valid proxies for cognitive empathy ability? Negligible meta-analytic relations with behavioral task performance. *Psychological Assessment, 31*, 1062–1072. https://doi.org/10.1037/pas0000732

Nader, A. M., Jelenic, P., & Soulieres, I. (2015). Discrepancy between WISC-III and WISC-IV cognitive profile in autism spectrum: What does it reveal about autistic cognition? *PLoS One, 10*, Article e0144645. https://doi.org/10.1371/journal.pone.0144645

Newbigin, A., Uljarevic, M., Vivanti, G., & Dissanayake, C. (2016). Brief report: Empathic responsiveness of high functioning children with autism to expressed and anticipated distress. *Journal of Autism and Developmental Disorders, 46*, 3338–3343. https://doi.org/10.1007/s10803-016-2862-x

Oakley, B. F. M., Brewer, R., Bird, G., & Catmur, C. (2016). Theory of mind is not theory of emotion: A cautionary note on the Reading the Mind in the Eyes Test. *Journal of Abnormal Psychology, 125*, 818–823. https://doi.org/10.1037/abn0000182

Oliver, L. D., Moxon-Emre, I., Hawco, C., Dickie, E. W., Dakli, A., Lyon, R. E., … Ameis, S. H. (2024). Task-based functional neural correlates of social cognition across autism and schizophrenia spectrum disorders. *Molecular Autism, 15*. https://doi.org/10.1186/s13229-024-00615-3

Page, M. J., Moher, D., Bossuyt, P. M., Boutron, I., Hoffmann, T. C., Mulrow, C. D., … McKenzie, J. E. (2021). PRISMA 2020 explanation and elaboration: Updated guidance and exemplars for reporting systematic reviews. *British Medical Journal, 372*. https://doi.org/10.1136/bmj.n160

Paulhus, D. L., & Vazire, S. (2007). The self-report method. In R. W. Robins, R. C. Fraley, & R. F. Krueger (Eds.), *Handbook of research methods in personality psychology* (pp. 224–239). New York: Guilford Press.

Paz, Y., Roth-Hanania, R., Gabis, L. V., Orlitsky, T., Zilka-Cohen, N., Zahn-Waxler, C., & Davidov, M. (2024). Predicting autism from young infants' empathic responding: A prospective study. *Development and Psychopathology*. https://doi.org/10.1017/S0954579424001226

Phan, K., Tian, D. H., Cao, C., Black, D., & Yan, T. D. (2015). Systematic review and meta-analysis: Techniques and a guide for the academic surgeon. *Annals of Cardiothoracic Surgery, 4*, 112–122. https://doi.org/10.3978/j.issn.2225-319X.2015.02.04

Polanin, J. R., & Terzian, M. (2019). A data-sharing agreement helps to increase researchers' willingness to share primary data: Results from a randomized controlled trial. *Journal of Clinical Epidemiology, 106*, 60–69. https://doi.org/10.1016/j.jclinepi.2018.10.006

Preston, S. D., & de Waal, F. B. (2002). Empathy: Its ultimate and proximate bases. *Behavioral and Brain Sciences, 25*, 1–20. https://doi.org/10.1017/s0140525x02000018

Pustejovsky, J. E., & Rodgers, M. A. (2019). Testing for funnel plot asymmetry of standardized mean differences. *Research Synthesis Methods, 10*, 57–71. https://doi.org/10.1002/jrsm.1332

R Core Team. (2021). *R: A language and environment for statistical computing*. Vienna, Austria: R Foundation for Statistical Computing. https://www.R-project.org/.

Regenbogen, C., Schneider, D. A., Finkelmeyer, A., Kohn, N., Derntl, B., Kellermann, T., … Habel, U. (2012). The differential contribution of facial expressions, prosody, and speech content to empathy. *Cognition & Emotion, 26*, 995–1014. https://doi.org/10.1080/02699931.2011.631296

Reniers, R. L., Corcoran, R., Drake, R., Shryane, N. M., & Vollm, B. A. (2011). The QCAE: A questionnaire of cognitive and affective empathy. *Journal of Personality Assessment, 93*, 84–95. https://doi.org/10.1080/00223891.2010.528484

Rieffe, C., O'Connor, R., Bülow, A., Willems, D., Hull, L., Sedgewick, F., Stockmann, L., & Blijd-Hoogewys, E. (2021). Quantity and quality of empathic responding by autistic and non-autistic adolescent girls and boys. *Autism, 25*, 199–209. https://doi.org/10.1177/1362361320956422

Rochat, M. J. (2023). Sex and gender differences in the development of empathy. *Journal of Neuroscience Research, 101*, 718–729. https://doi.org/10.1002/jnr.25009

Rødgaard, E. M., Jensen, K., & Mottron, L. (2019). An opposite pattern of cognitive performance in autistic individuals with and without alexithymia. *Journal of Abnormal Psychology, 128*, 735–737. https://doi.org/10.1037/abn0000408

Rødgaard, E. M., Jensen, K., Vergnes, J. N., Soulieres, I., & Mottron, L. (2019). Temporal changes in effect sizes of studies comparing individuals with and without autism: A meta-analysis. *JAMA Psychiatry*. https://doi.org/10.1001/jamapsychiatry.2019.1956

Rodgers, M. A., & Pustejovsky, J. E. (2021). Evaluating meta-analytic methods to detect selective reporting in the presence of dependent effect sizes. *Psychological Methods, 26*, 141–160. https://doi.org/10.1037/met0000300

Santiesteban, I., Gibbard, C., Drucks, H., Clayton, N., Banissy, M. J., & Bird, G. (2021). Individuals with autism share others' emotions: Evidence from the continuous affective rating and empathic responses (CARER) task. *Journal of Autism and Developmental Disorders, 51*, 391–404. https://doi.org/10.1007/s10803-020-04535-y

Scambler, D. J., Hepburn, S., Rutherford, M. D., Wehner, E. A., & Rogers, S. J. (2007). Emotional responsivity in children with autism, children with other developmental disabilities, and children with typical development. *Journal of Autism and Developmental Disorders, 37*, 553–563. https://doi.org/10.1007/s10803-006-0186-y

Scheeren, A. M., Koot, H. M., Mundy, P. C., Mous, L., & Begeer, S. (2013). Empathic responsiveness of children and adolescents with high-functioning autism spectrum disorder. *Autism Research, 6*, 362–371. https://doi.org/10.1002/aur.1299

Schneider, K., Regenbogen, C., Pauly, K. D., Gossen, A., Schneider, D. A., Mevissen, L., … Schneider, F. (2013). Evidence for gender-specific endophenotypes in high-functioning autism spectrum disorder during empathy. *Autism Research, 6*, 506–521. https://doi.org/10.1002/aur.1310

Schulte-Rüther, M., Greimel, E., Markowitsch, H. J., Kamp-Becker, I., Remschmidt, H., Fink, G. R., & Piefke, M. (2011). Dysfunctions in brain networks supporting empathy: An fMRI study in adults with autism spectrum disorders. *Social Neuroscience, 6*, 1–21. https://doi.org/10.1080/17470911003708032

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

Sheppard, E., Pillai, D., Wong, G. T., Ropar, D., & Mitchell, P. (2016). How easy is it to read the minds of people with autism spectrum disorder? *Journal of Autism and Developmental Disorders, 46*, 1247–1254. https://doi.org/10.1007/s10803-015-2662-8

Sher, M. A., & Oliver, C. (2022). Assessment of self-report response bias in high functioning autistic people. *Psychiatry, Psychology and Law*. https://doi.org/10.1080/13218719.2021.2006097

Shi, L., & Lin, L. (2019). The trim-and-fill method for publication bias: Practical guidelines and recommendations based on a large database of meta-analyses. *Medicine, 98*, Article e15987. https://doi.org/10.1097/MD.0000000000015987

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, 63*, 796–807. https://doi.org/10.1111/j.1467-8624.1992.tb01662.x

Silke, C., Brady, B., Boylan, C., & Dolan, P. (2018). Factors influencing the development of empathy and pro-social behaviour among adolescents: A systematic review. *Children and Youth Services Review, 94*, 421–436. https://doi.org/10.1016/j.childyouth.2018.07.027

Singer, T. (2006). The neuronal basis and ontogeny of empathy and mind reading: Review of literature and implications for future research. *Neuroscience and Biobehavioral Reviews, 30*, 855–863. https://doi.org/10.1016/j.neubiorev.2006.06.011

Singer, T., & Lamm, C. (2009). The social neuroscience of empathy. *Annals of the New York Academy of Sciences, 1156*, 81–96. https://doi.org/10.1111/j.1749-6632.2009.04418.x

Singer, T., Seymour, B., O'Doherty, J., Kaube, H., Dolan, R. J., & Frith, C. D. (2004). Empathy for pain involves the affective but not sensory components of pain. *Science, 303*, 1157–1162. https://doi.org/10.1126/science.1093535

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

Smith, A. (2009). The empathy imbalance hypothesis of autism: A theoretical approach to cognitive and emotional empathy in autistic development. *The Psychological Record, 59*, 273–294. https://doi.org/10.1007/BF03395663

Song, Y., Nie, T., Shi, W., Zhao, X., & Yang, Y. (2019). Empathy impairment in individuals with autism spectrum conditions from a multidimensional perspective: A meta-analysis. *Frontiers in Psychology, 10*. https://doi.org/10.3389/fpsyg.2019.01902

Stanley, T. D., Carter, E. C., & Doucouliagos, H. (2018). What meta-analyses reveal about the replicability of psychological research. *Psychological Bulletin, 144*, 1325–1346. https://doi.org/10.1037/bul0000169

Stroth, S., Paye, L., Kamp-Becker, I., Wermter, A. K., Krach, S., Paulus, F. M., & Muller-Pinzler, L. (2019). Empathy in females with autism spectrum disorder. *Frontiers in Psychiatry, 10*. https://doi.org/10.3389/fpsyt.2019.00428

Sun, R. W., & Cheung, S. F. (2020). The influence of nonnormality from primary studies on the standardized mean difference in meta-analysis. *Behavior Research Methods, 52*, 1552–1567. https://doi.org/10.3758/s13428-019-01334-x

Sunahara, C. S., Rosenfield, D., Alvi, T., Wallmark, Z., Lee, J., Fulford, D., & Tabak, B. A. (2022). Revisiting the association between self-reported empathy and behavioral assessments of social cognition. *Journal of Experimental Psychology: General, 151*, 3304–3322. https://doi.org/10.1037/xge0001226

Thaler, H., Skewes, J. C., Gebauer, L., Christensen, P., Prkachin, K. M., & Elmholdt, E.-M. J. (2018). Typical pain experience but underestimation of others' pain: Emotion perception in self and others in autism spectrum disorder. *Autism, 22*, 751–762. https://doi.org/10.1177/1362361317701269

Tkalcec, A., Bierlein, M., Seeger-Schneider, G., Walitza, S., Jenny, B., Menks, W. M., … Cubillo, A. (2023). Empathy deficits, callous-unemotional traits and structural underpinnings in autism spectrum disorder and conduct disorder youth. *Autism Research, 16*, 1946–1962. https://doi.org/10.1002/aur.2993

Tsang, T., Gillespie-Lynch, K., & Hutman, T. (2016). Theory of mind indexes the broader autism phenotype in siblings of children with autism at school age. *Autism Research and Treatment, 2016*. https://doi.org/10.1155/2016/6309189

Vachon, D. D., & Lynam, D. R. (2016). Fixing the problem with empathy: Development and validation of the affective and cognitive measure of empathy. *Assessment, 23*, 135–149. https://doi.org/10.1177/1073191114567941

Van den Noortgate, W., López-López, J. A., Marín-Martínez, F., & Sánchez-Meca, J. (2013). Three-level meta-analysis of dependent effect sizes. *Behavior Research Methods, 45*, 576–594. https://doi.org/10.3758/s13428-012-0261-6

Van den Noortgate, W., López-López, J. A., Marín-Martínez, F., & Sánchez-Meca, J. (2015). Meta-analysis of multiple outcomes: A multilevel approach. *Behavior Research Methods, 47*, 1274–1294. https://doi.org/10.3758/s13428-014-0527-2

Viechtbauer, W. (2010). Conducting meta-analyses in R with the metafor package. *Journal of Statistical Software, 36*, 1–48. https://doi.org/10.18637/jss.v036.i03

Viechtbauer, W., & Cheung, M. W. (2010). Outlier and influence diagnostics for meta-analysis. *Research Synthesis Methods, 1*, 112–125. https://doi.org/10.1002/jrsm.11

Vilas, S. P., Reniers, R., & Ludlow, A. K. (2021). An investigation of behavioural and self-reported cognitive empathy deficits in adolescents with autism spectrum disorders and adolescents with behavioural difficulties. *Frontiers in Psychiatry, 12*. https://doi.org/10.3389/fpsyt.2021.717877

de Waal, F. B. M., & Preston, S. D. (2017). Mammalian empathy: Behavioural manifestations and neural basis. *Nature Reviews. Neuroscience, 18*, 498–509. https://doi.org/10.1038/nrn.2017.72

Wakabayashi, A., Baron-Cohen, S., Uchiyama, T., Yoshida, Y., Kuroda, M., & Wheelwright, S. (2007). Empathizing and systemizing in adults with and without autism spectrum conditions: Cross-cultural stability. *Journal of Autism and Developmental Disorders, 37*, 1823–1832. https://doi.org/10.1007/s10803-006-0316-6

Walter, H. (2012). Social cognitive neuroscience of empathy: Concepts, circuits, and genes. *Emotion Review, 4*, 9–17. https://doi.org/10.1177/1754073911421379

Zadok, E., Gordon, I., Navon, R., Rabin, S. J., & Golan, O. (2021). Shifts in behavioral synchrony in response to an interaction partner's distress in adolescents with and without ASD. *Journal of Autism and Developmental Disorders*. https://doi.org/10.1007/s10803-021-05307-y

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

Zaki, J., & Ochsner, K. N. (2012). The neuroscience of empathy: Progress, pitfalls and promise. *Nature Neuroscience, 15*, 675–680. https://doi.org/10.1038/nn.3085

Zaki, J., Weber, J., Bolger, N., & Ochsner, K. (2009). The neural bases of empathic accuracy. *Proceedings of the National Academy of Sciences of the United States of America, 106*, 11382–11387. https://doi.org/10.1073/pnas.0902666106