

Developmental Review 57 (2020) 100925



Contents lists available at ScienceDirect

Developmental Review

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

Metaphor processing in autism: A systematic review and meta-analysis

Kinga Morsanyi^{a,b,*}, Dušan Stamenković^c, Keith J. Holyoak^d^a School of Psychology, Queen's University Belfast, Northern Ireland, United Kingdom^b Centre for Mathematical Cognition, Loughborough University, United Kingdom^c Faculty of Philosophy & Center for Cognitive Sciences, University of Niš, Serbia^d Department of Psychology, University of California, Los Angeles, USA

ARTICLE INFO

Keywords:

Autism spectrum disorder
Conceptual integration
Figurative language
Meta-analysis
Metaphor comprehension
Systematic review
Verbal abilities

ABSTRACT

Impairments related to figurative language understanding have been considered to be one of the diagnostic and defining features of autism. Metaphor comprehension and production in autism spectrum disorder (ASD) as compared to typically developing (TD) individuals have been investigated for around thirty years, generally showing an overall advantage for TD groups. We present a preregistered systematic review and meta-analysis including a total of 15 studies that fulfilled our set of inclusion criteria (notably, ASD and TD groups matched in chronological age and verbal- or full-scale IQ). Along with accuracy, we also analyzed group differences in reaction time in the studies that reported them. The results revealed a medium-to-large group difference favoring TD over ASD groups based on accuracy measures, as well as a similar overall advantage for TD groups based on reaction times. There was reliable heterogeneity in effect sizes for group differences in accuracy, which was mostly explained by the effect of verbal intelligence, with differences in metaphor processing being smaller for participants with better verbal skills. Some of the variation in effect sizes may also be attributed to differences in types of metaphor processing tasks. We also evaluated the quality of the studies included in the meta-analysis, and the evidence relating to the potential presence of publication bias.

Introduction

Continuing controversy surrounds the nature of the social, cognitive and linguistic differences between people diagnosed with Autistic Spectrum Disorder (ASD) in comparison with the typically-developing (TD) population. People with ASD are often characterized as having difficulty grasping the pragmatics of social situations, lacking in cognitive flexibility, oriented toward detailed perceptual features, weak in central coherence, and impaired in language processing (e.g., Frith, 2003; Happé & Frith, 2006; Pellicano, Maybery, Durkin, & Maley, 2006; Rajendran & Mitchell, 2007). At the same time, empirical findings suggest that some high-level cognitive abilities, notably analogical reasoning, may be spared in autism (e.g., Scott & Baron-Cohen, 1996; Dawson, Soulières, Gernsbacher, & Mottron, 2007; Morsanyi & Holyoak, 2010), or even constitute an area of relative cognitive strength (for a recent meta-analysis see Morsanyi, Stamenković, & Holyoak, 2020). In addition, there is evidence that people with ASD exhibit reduced susceptibility to some common reasoning errors and biases, such as the conjunction fallacy (Morsanyi, Handley, & Evans, 2010), the framing effect (De Martino, Harrison, Knafo, Bird, & Dolan, 2008), and the jumping-to-conclusion bias (Brosnan, Chapman, & Ashwin, 2014).

* Corresponding author at: Schofield Building, Loughborough University, Ashby Road, Loughborough LE11 3TT, United Kingdom.
E-mail address: K.E.Morsanyi@lboro.ac.uk (K. Morsanyi).

Metaphor processing

Impairments of figurative language understanding have been considered to be a defining feature of autism, and one area of particular interest is the ability to comprehend and produce metaphors. Metaphor is the use of language to describe one thing in terms of something else that is conceptually very different, as in “The job fair was a circus.” The syntactic and conceptual forms of metaphors can vary enormously. As a more complex example, consider “His fine wit makes such a wound, the knife is lost in it” (a quote from the poet Percy Bysshe Shelley). Here the conventional notion of “barbed” words that “cut” their receiver is extended to suggest how the impact of an insult can be so rapid and so severe that the victim is unable to grasp exactly what happened.

Metaphor processing, which appears to depend on both language and thinking abilities, improves over the course of typical cognitive development (Rundblad & Annaz, 2010). Early signs of metaphor appreciation appear at least by age four (Gentner, 1988; Vosniadou, 1987), with development continuing through childhood (Pouscoulous, 2011, 2014; Vosniadou, Ortony, Reynolds, & Wilson, 1984), adolescence, and into adulthood (Nippold, 2016). There appears to be a marked improvement in metaphor processing around the age of 10 (Lecce, Ronchi, Del Sette, Bischetti, & Bambini, 2019; Winner, Rosenstiel, & Gardner, 1976). Lecce et al. (2019) found that abilities involving theory of mind (ToM) were related to metaphor processing among children under the age of ten, but this link was not present in the case of older children.

Apart from a potential influence of theory of mind, metaphor processing relies heavily on linguistic skills, including lexical knowledge (especially in the case of conventional metaphors; Pouscoulous, 2014) and pragmatic inference (Recanati, 2004; Sperber & Wilson, 2012; Wilson & Carston, 2006). It is known that deficits in general language processing, even in the absence of ASD, are associated with deficits in metaphor processing (e.g., Highnam, Wegmann, & Woods, 1999). Given that ASD is often associated with varieties of linguistic impairment, as well as impairment of theory of mind, metaphor comprehension is likely to also be compromised across the ASD population as a whole. Indeed, the latest edition of the Diagnostic and Statistical Manual of the American Psychiatric Association (DSM-5; 2013) lists difficulties understanding nonliteral and ambiguous meanings of language (as in metaphors) among the core characteristics required for an ASD diagnosis. Nevertheless, it has been argued that pragmatic impairments in autism are neither global nor uniform (Geurts, Kissine, & van Tiel, 2020), and it is unclear whether people with ASD have problems with metaphor processing that go beyond more general language impairment (for critical assessments see Brock, Norbury, Einav & Nation, 2008; Gernsbacher & Pripas-Kapit, 2012; Norbury, 2005).

A number of competing theories of typical metaphor processing have been proposed (for recent reviews see Holyoak, 2019; Holyoak & Stamenković, 2018). Older theories (e.g., Searle, 1983) claimed that language comprehension proceeds in a “literal first” manner, with metaphorical processing being evoked only if the literal meaning fails to make sense in context. However, there is strong evidence that metaphor processing begins prior to extraction of the literal meaning (e.g., Glucksberg, Gildea, & Bookin, 1982). Indeed, some theories claim that metaphorical and literal language are so deeply intertwined as to be virtually identical in their processing requirements (e.g., Lakoff & Johnson, 1980). To the extent that metaphor processing fully overlaps with processing of literal language, it would appear unlikely that ASD (or any other condition) would have a selective impact on metaphor processing *per se*.

Other theories have claimed that specific cognitive processes are involved in metaphor processing. Though these processes may also be evoked at least sometimes in processing literal language, as well as in nonlinguistic tasks, they are believed to be especially prominent in metaphor processing. A classical idea that traces to Aristotle is that metaphor is based on analogical reasoning—the ability to find and exploit similarities based on *relations* among entities, rather than solely on the entities themselves (for a review of research on analogy see Holyoak, 2012). In modern psychology, the analogy hypothesis was developed by Tourangeau and Sternberg (1981) and Gentner and Clement (1988). For example, in “The job fair was a circus,” the job fair is a relatively unfamiliar *target* that is being systematically compared to a circus, which serves as the more familiar *source*. A weaker version of the hypothesis claims that analogy is used for novel or unfamiliar metaphors, but not for conventional or familiar ones (Bowdle & Gentner, 2005). As noted above, current evidence indicates that analogical reasoning (in nonverbal tasks) is spared in ASD (Morsanyi, Stamenković, & Holyoak, 2020). This meta-analytic study found that autistic individuals even showed superior performance (with a medium effect size) when compared to age and Wechsler IQ-matched controls in the case of perceptual analogies, notably Raven’s Progressive Matrices. Thus, if analogical reasoning were the sole critical mechanism for metaphor processing, selective impairment of metaphor processing in ASD would not be expected.

An alternative theoretical approach claims that metaphors (at least nominal metaphors, those in the form “NOUN is NOUN”) are interpreted as category statements (Glucksberg & Keysar, 1990). For example, in “The job fair was a circus,” the source may be interpreted not as a “real” circus with clowns and animals, but as something like the abstract category of “events that involve many complicated activities happening at once.” Whereas the analogy view assumes that processing a nominal metaphor involves finding systematic correspondences between elements that comprise the source and target, the categorization view assumes that metaphor comprehension operates by categorizing the target as a member of an abstract category derived from the source. A more general version of the categorization hypothesis proposes that metaphor comprehension involves a process of *conceptual combination*, in which the semantic representations of individual words are integrated based on their semantic features (Kintsch, 2000). This approach views the process underlying metaphor comprehension as a relatively complex version of a mechanism for integrating word meanings that contributes to literal comprehension (e.g., Wieniewski, 1997). If a disorder such as ASD were to impair the ability to

happen automatically in the case of TD individuals, for whom this tendency is so strong that it often leads to reasoning biases and memory illusions. For example, when people are presented with a list of closely related words (e.g., bed, rest, tired, dream, blanket), and then are asked to recall the presented words, they commonly also list non-presented but closely-related words (e.g., sleep). This tendency appears to be reduced in ASD (e.g., Wojcik et al., 2018). More broadly, a recent meta-analysis found that ASD individuals display decreased suggestibility and are less likely to develop false memories than TD controls (Griego, Datzman, Estrada & Middlebrook, 2019). A similar lack of contextual processing may explain reduced susceptibility to some reasoning biases in autism, such as the conjunction fallacy (Morsanyi, Handley, & Evans, 2010), in which TD individuals tend to base their judgments on the gist of extracted meaning, rather than on processing of verbatim detail. Overall, these findings indicate a reduced tendency for semantic integration and conceptual combination in ASD, which may be critical mechanisms for metaphor processing.

In the present paper, we address the question of whether ASD is associated with a selective deficit in metaphor processing by performing a systematic literature review and meta-analysis.

Reviews and theoretical discussions of metaphor processing in ASD

The first accounts of metaphor processing in autism attributed pragmatic difficulties to problems with mind reading (Baron-Cohen, Tager-Flusberg & Cohen, 2000; Happé, 1993). However, this interpretation has been challenged by several theorists who proposed that a more general language impairment (in particular, problems with semantic knowledge) and structural language skills explain the difficulties with metaphor processing, and with figurative language in general (Brock et al., 2008; Gernsbacher & Pripas-Kapit, 2012; Geurts, Kissine, & van Tiel, 2020; Norbury, 2005).

Kalandadze, Norbury, Nærland, and Næss (2018) performed a meta-analytic review to investigate this hypothesis. Their review covered a wide variety of figurative language. In addition to metaphor, they reviewed studies of verbal irony, sarcasm, hyperbole, metonymy, idioms, proverbs, and humor, with some studies reporting results that covered various combinations of these. Although these phenomena may well be interrelated, they are extremely heterogeneous and may not form a unitary class. Kalandadze et al. (2018) focused on group comparisons based on a combined measure across all forms of figurative language. Overall, the meta-analysis indicated that individuals with ASD exhibit moderately poorer figurative language comprehension skills in comparison to TD controls, and that this deficit covaries with social communication problems in ASD. However, studies in which the ASD and TD groups were matched on verbal ability yielded nonsignificant differences. At a more fine-grained level, metaphors were generally more difficult to comprehend for ASD individuals relative to TD controls than were instances of sarcasm and irony. Kalandadze et al. concluded that the figurative language deficit in ASD is related to broader issues involving language skills. Furthermore, these authors proposed that impairments in core language abilities might explain problems with social communication in ASD, rather than the other way around.

In a critical overview of experimental and clinical research related to problems in figurative language processing in autism, Vulchanova, Saldaña, Chauboun, and Vulchanov (2015) observed that problems with extended and figurative language persist even in the cases of linguistically-talented individuals with Asperger's syndrome. Vulchanova et al. (2015) noted that a wide range of cognitive and linguistic abilities may play important roles in processing figurative language, including syntactic and semantic knowledge, vocabulary size, conceptual knowledge base, skill in drawing inferences and performing information integration, mentalizing and understanding intentions, and the ability to suppress irrelevant information. They hypothesized that impairment in any one of these components can lead to problems in figurative language processing. Although in the case of high-functioning autistic individuals most of these competencies are intact, and these individuals can also pass first- and second-order ToM tasks, they still experience problems with information integration. In particular, they have a reduced ability to evaluate the plausibility and relevance of different pieces of information. Moreover, they struggle with combining information from different sensory modalities.

In another recent review of figurative language in general, Saban-Bezalel and Mashal (2018) concluded that ASD deficits in metaphor processing are present for children and adolescents, but they may be absent in adulthood (although performance depends on the type of metaphor task and the format of required responses). These authors reviewed a study by Kasirer and Mashal (2014) in which autistic and TD adults were tested on their ability to comprehend and generate metaphors and similes. Kasirer and Mashal found no group differences in comprehension abilities. Moreover, autistic participants generated a larger number of metaphors and similes, and their responses were judged to be more creative by independent evaluators (e.g., "Feeling successful is like seeing the view from the mountaintop."). Saban-Bezalel and Mashal suggested that the ability of autistic individuals to generate unconventional responses may reflect a reduced ability to take into account the addressee's point of view.

Another recent meta-analytic study (Kalandadze, Bambini, & Naess, 2019) focused solely on metaphor processing in ASD, with a particular interest in the potential effects of task properties on performance. This study obtained a medium effect size for group differences (with a reduced ability in the ASD group to process metaphors). Due to the relatively small number of studies, and because the tasks tended to differ on many dimensions, the study was unable to quantitatively analyze the effects of task properties. Nevertheless, Kalandadze et al. (2019) noted that tasks requiring verbal explanations appeared to be the most challenging for autistic participants, whereas decisions about meaningfulness appeared to be the easiest. These authors proposed that tasks requiring the integration of multiple modalities may pose problems for autistic individuals (although this observation was based on the results of only two studies). Notably, this review and meta-analysis did not address some important controversies in the literature, such as the

Kapit, 2012; Geurts et al., 2020; Kalandadze et al., 2018; Norbury, 2005) have concluded that group differences disappear when the ASD and TD groups are appropriately matched on verbal ability. Metaphor is a particularly challenging form of figurative language, as it requires finding and exploiting connections between relatively remote semantic domains (e.g., verbal and physical assault in the example from Shelley). This is a distinctive aspect of metaphor processing, which is not shared with other forms of nonliteral language. As noted above, Kalandadze et al. (2018) found that metaphor processing was more difficult for participants with ASD than other forms of non-literal language, such as sarcasm and irony.

In their review, Vulchanova et al. (2015) concluded that differences in metaphor processing were still present in the case of high-functioning autistic individuals, due to their difficulties with information integration. Moreover, Vulchanova et al. noted that a range of cognitive and linguistic abilities play a role in metaphor processing, and impairment in any of these abilities can lead to problems with figurative language processing. Although not explicitly stated in their review, these observations suggest that difficulties in metaphor processing might be exacerbated in the case of autistic participants with a learning disability, because they might be impaired in multiple cognitive components necessary for metaphor processing.

Saban-Bezalel and Mashal (2018) proposed that impairments in metaphor processing might be present in childhood, but that this group difference is no longer present for high-functioning adults with ASD. This account suggests a potential developmental delay in metaphor processing, with group differences diminishing with age (at least for high-functioning individuals). Kalandadze et al. (2019) identified some task properties (e.g., asking participants to provide a verbal explanation) that may affect the size of group differences in metaphor processing between ASD and TD individuals. Overall, some fundamental questions about metaphor processing in autism remain to be answered. Our aim in the present review is to address these issues based on currently available evidence.

Aims and scope of the present review

In the present review, similar to Kalandadze et al. (2019), we focus specifically on metaphor processing in autism. In order to more cleanly identify studies that could shed light on a possible selective impairment in metaphor processing, as opposed to a more general cognitive or linguistic impairment, our meta-analysis sought studies in which both chronological age and IQ (ideally, verbal IQ) were equated across the ASD group and a TD control group. This way, we were able to address the question of whether group differences still exist when the groups are matched on intellectual ability (and more specifically, verbal intelligence). In addition to performing a meta-analysis of the effect size of group differences in metaphor processing, we also performed a meta-regression to determine whether the size of group differences in metaphor processing is related to participants' level of verbal intelligence. Specifically, it is possible that autistic individuals with more profound language impairment find metaphor processing particularly challenging, whereas group differences may be small or absent in the case of autistic individuals with average or high verbal intelligence.

We also addressed the question of whether there was a developmental delay in metaphor processing by performing a meta-regression with chronological age as a predictor of the size of group differences in metaphor processing. It is possible that (high functioning) autistic individuals develop the ability to process metaphors, but acquiring this skill is delayed relative to their TD peers. If this is the case, we could expect smaller or non-existent group differences in the case of older participants. We also considered the possibility that some task properties may be linked to the size of group differences (see Kalandadze et al., 2019, for a detailed discussion). Finally, it is possible that this literature is affected by a publication bias, because it is generally easier to publish studies that report positive results (i.e., a significant group difference in metaphor processing), and also because a group difference in metaphor processing could be expected based on the official diagnostic criteria of ASD. Thus, we also performed analyses related to the presence of a publication bias. Our study was preregistered in the International Register of Systematic Reviews, PROSPERO (available at: https://www.crd.york.ac.uk/PROSPERO/display_record.php?RecordID=116248).

Method

Search strategy

In designing our systematic review and reporting the results of the meta-analysis, we followed the *Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA)* guidelines. We searched the following electronic databases: *PsycARTICLES*, *PsycINFO*, *ERIC*, *Linguistics and Language Behavior Abstracts (LLBA)*, *ProQuest Dissertations & Theses Global* (all five searched using the *ProQuest* engine), *Web of Science* and *Scopus*, for all studies published up to July 2020. We used the following combinations of search terms: *autis* OR ASD OR Asperger** crossed with *metaphor**. The target fields in the searches included titles, abstracts, keywords, topics, subjects and indexing. No restrictions in terms of the publication year were applied.

Notably, our search also covered the gray literature (i.e., studies that are not published in peer-reviewed journals, such as conference proceedings and dissertations). This way, we aimed to minimize the effect of publication bias (i.e., an overrepresentation of studies that report positive findings) on our meta-analysis.¹ It is important to note that small studies with non-significant findings

are the most likely to remain unpublished (cf. [Conn, Valentine, Cooper & Rantz, 2003](#)). Leaving out these studies from meta-analyses is undesirable because one of the main advantages of meta-analyses is that results can be summarized across small samples (i.e., it is possible that a statistically significant effect emerges when pooling smaller studies that yielded non-significant results). Another important consideration is that sample sizes and effect sizes tend to be negatively correlated in meta-analyses (e.g., [Levine, Asada & Carpenter, 2009](#)). That is, the reporting of small studies is biased in that they are more likely to be included in meta-analyses when they yielded significant results.

Study inclusion criteria

We selected articles for the meta-analysis on the basis of the following predetermined criteria:

- (1) The paper reported the results of an original research study including a metaphor comprehension or production task, in which scores relating to accuracy on the relevant task were reported separately from other measures used in the study. Studies were excluded if insufficient data were available to calculate effects sizes and relevant data could not be obtained from the author(s).
- (2) Participants had to be diagnosed with ASD by experienced clinicians using the Diagnostic and Statistical Manual of Mental Disorders (DSM) or International Classification of Diseases (ICD) diagnostic criteria. This included the previously separate diagnostic categories of autistic disorder, Asperger's syndrome and pervasive developmental disorder (not otherwise specified).
- (3) The study had to include a TD comparison group, matched to the ASD group on chronological age and either full-scale or verbal IQ.

When we planned the study and performed our initial literature searches, we were not aware of the meta-analytic study on metaphor processing in autism performed by [Kalandadze et al. \(2019\)](#). The overlap between the two studies is not extensive. [Kalandadze et al. \(2019\)](#) included 14 studies in their analysis. Our inclusion criteria filtered out nine of these studies, while including eleven additional studies (presented in ten papers), with only five studies overlapping between the two meta-analyses.

Screening process

The search was conducted by two independent raters, D.S. and K.M., using three major databases and search engines, and any disagreements were resolved by discussions among all three authors. The first step was to remove any duplicates by comparing the three lists of results (coming from *ProQuest*, *Scopus* and *Web of Science*). The next step, performed by D.S. and K.M., involved reading the titles and the abstracts of each study. When these provided insufficient information, the full-text article was reviewed. The inclusion criteria were applied to select from the full-text articles the final papers to be included in the meta-analysis. The final selection was performed by K.M. and K.H., who also extracted the data from the papers for the purpose of the meta-analytic analyses. Any disagreements were resolved by discussion. Further details on the number of papers included in each step, and the reasons for exclusion, are presented in the PRISMA flow diagram ([Moher et al., 2009](#)) displayed in [Fig. 1](#).

Coding

Study characteristics (title, authors and publication year) were coded for descriptive purposes. For statistical analyses, we coded the number of ASD and TD participants, and the inferential statistics reported in the papers. The statistics of primary interest were means and standard deviations relating to performance of each group on the metaphor tasks. For studies that included neuroimaging data, only the behavioral results were coded. When a study included multiple tasks relating to metaphor processing, or the results were broken down by type of metaphor task (for example, conventional and novel metaphors), these results were combined into a single measure for the study, because computing effect sizes multiple times based on data from the same sample can distort the overall results ([Borenstein, Hedges, Higgins, & Rothstein, 2009](#)). Type of task (including the presentation format and required response) was coded. We also coded the average age and the IQ of the participants in the ASD and TD samples, and the strategy used for matching the groups on intellectual ability (i.e., verbal or full-scale IQ). Matching strategy was not used as a variable in our analyses, because there was only a single study that matched the samples on full-scale IQ.

Meta-analytic procedures and analysis

All statistical analyses were conducted using the Comprehensive Meta-Analysis (CMA) software, version 3 (Biostat). A 95% confidence interval was computed for each effect size to indicate if it was statistically different from zero (i.e., if the confidence interval does not include zero, the effect is considered significant).

Based on the original data from each study, effect sizes were computed using Hedges' g (a variation of Cohen's d that corrects for

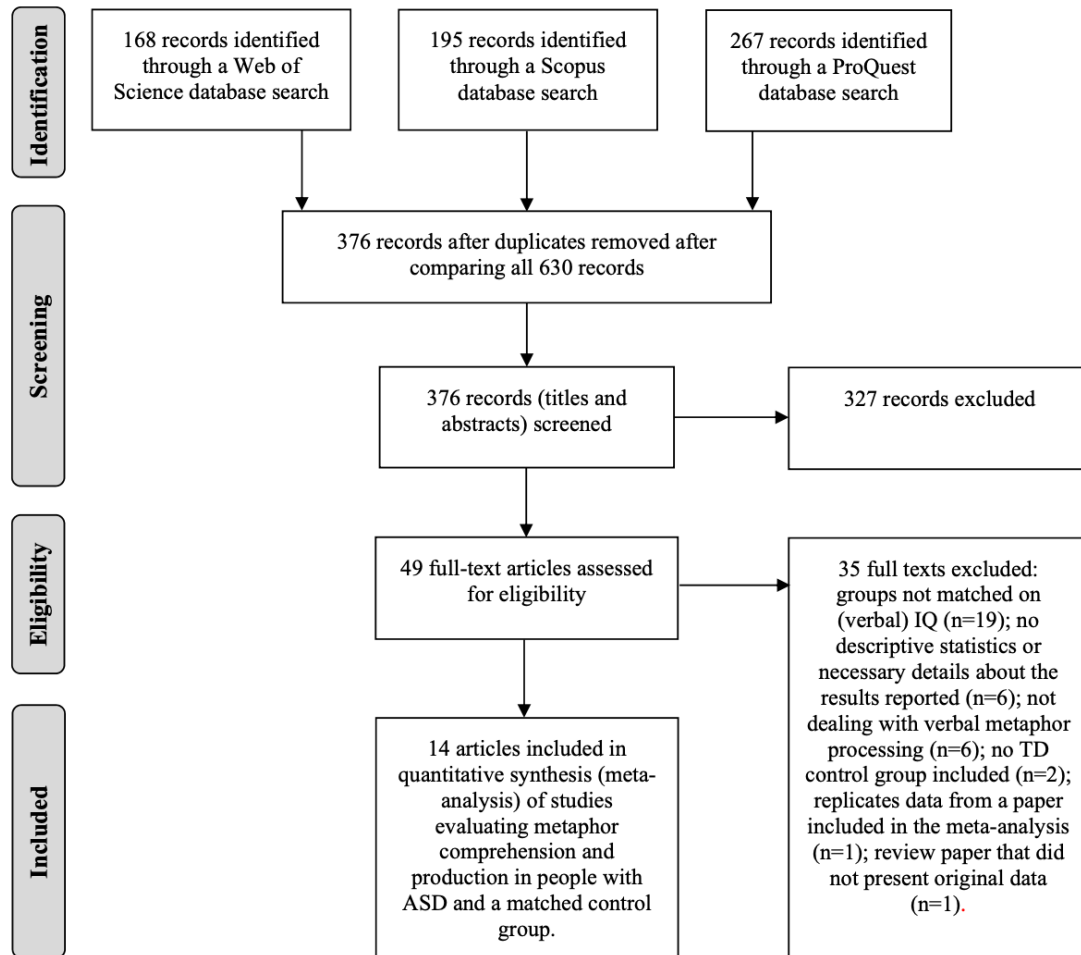


Fig. 1. The PRISMA flow diagram describing the search process, the number of papers included in each step, and the reasons for exclusion at the full-text stage.

biases due to small sample sizes; Hedges, 1981; Hedges & Olkin, 1985). The Hedges' g value was defined as positive when TD individuals had the higher group mean (and vice versa for negative values). The overall effect size was estimated by calculating a weighted average of individual effect sizes, based on a random effects model that assumes between-study variations in effect sizes result not only from random error, but also from systematic effects of some variables that are likely to vary from study to study (Borenstein et al., 2009).² The assumption that effect sizes are heterogeneous allows the possibility that factors beyond an ASD diagnosis impact effect sizes. Heterogeneity of effect sizes was statistically tested using Cochran's Q -statistic. We also report the I^2 statistic, which expresses the percentage of variation in effect sizes across studies that is due to systematic effects of study variables, rather than chance (Higgins & Thompson, 2002; Higgins, Thompson, Deeks, & Altman, 2003). Meta-regression analyses, using random-effects models, were carried out to test for the effects of possible moderator variables (i.e., the average age and IQ of ASD participants).

In order to identify and statistically evaluate the effect of publication bias on our overall estimate of effect size, we created a funnel plot, in which we plotted the effect sizes of group difference on the horizontal axis against the standard error of effect sizes (a sample-size dependent statistic) on the vertical axis. Studies with smaller standard errors appear toward the top of the graph and tend to cluster around the mean effect size; those with larger standard errors appear toward the bottom, and tend to be more dispersed. In the absence of publication bias, studies are expected to be symmetrically distributed on each side of the overall mean effect size (Cooper, Hedges, & Valentine, 2009). If publication bias is present, a higher concentration of studies is expected on one side of the mean toward the bottom of the plot. In the case of a random effects model, it can be difficult to interpret the funnel plot visually (Lau, Ioannidis, Terrin, Schmid, & Olkin, 2006). We therefore used the classic fail-Safe N statistic (Rosenthal, 1979) to statistically test for publication bias, although this result should also be treated with caution when a relatively small number of studies are included in the

analysis (Lau et al., 2006). The fail-safe N analysis investigates the stability of the findings of the meta-analysis by assessing the degree to which including additional studies with non-significant findings would change the overall result to non-significant (Long, 2001). The aim of this analysis is to try to tackle a potential file drawer problem such that non-significant findings are less likely to be published. As a rule of thumb, Rosenthal (1969) proposed that the results can be considered stable if the number of missing studies that would change the result to non-significant is above $5k + 10$, where k represents the number of studies in the meta-analysis. Thus, in the present case, with 15 studies, the critical number of missing studies would have to be above 85 for the results to be considered stable. As an additional check of publication bias, we also computed the correlation between effect sizes and sample sizes across the studies included in our meta-analysis. A significant negative correlation might be found in meta-analyses when small studies are only included if they yielded a significant result (e.g., Levine et al., 2009).

We have also assessed the quality of the included studies, using a similar procedure to Spain, Sin, Linder, McMahon, and Happé (2018), who based their approach on the quality assessment tool for quantitative studies developed by Thomas, Ciliska, Dobbins, and Micucci (2004), but omitting components that are only relevant for intervention studies. In addition, when we decided on the assessment criteria, we considered the fact that the studies included in our analysis all compared an atypical group to a TD group, which made it necessary for the groups to be matched on some important characteristics. We expected that in particular three aspects of the studies that were included in our meta-analysis would be critical in relation to the quality of the results: sample characteristics, quality of the metaphor task (i.e., the main outcome measure), and procedure and materials used for matching the groups on verbal or intellectual ability. With regard to *sample characteristics* we considered whether the study sought to match participants in terms of their baseline demographic characteristics (apart from age, whether the participants were matched on gender, socio-economic status, and educational background); whether the ASD and control samples were recruited in the same way (e.g., from the same schools or higher education settings, or from the broader community); and in studies where the ASD group was recruited from a clinical setting, whether they had similar demographics to controls. We also considered the potential consequences of exclusions and withdrawals. In terms of the *quality of the metaphor task*, we considered whether the authors took some steps to ensure that the metaphor task had good psychometric properties (e.g., by using a standardised instrument or piloting the task with an independent sample, checking the reliability of the task, and using a relatively large number of items). Regarding the *procedure and materials used for matching the groups on verbal or intellectual ability* we considered whether a standardised instrument was used to match participants on (verbal) IQ. In case other types of (verbal) ability measures were used, we considered whether there was a risk that the groups differed in basic language skills, such as vocabulary or the ability to follow instructions. Based on these considerations, each aspect was assigned a rating of strong, moderate or weak by two independent raters (K.M. and D.S.). Based on these evaluations, both raters assigned an overall rating to each study, also using intermediate categories (e.g., “moderate to strong”). The final rating for each study was agreed upon by both raters.

Results

A total of 15 studies (published in 14 papers), involving 351 ASD and 354 TD participants, were included in the analyses. The country and language of each study, the characteristics of the participants included in each study in the meta-analysis (number of participants in the ASD and TD groups, and their mean chronological age and verbal- or full-scale IQ) and the type of metaphor task (s) used in each study, are listed in Table 1. The studies are listed in the table in rank order of effect size (Hedge's g) for the group difference in performance on the metaphor task, from negative (i.e., ASD group performs better) to positive (TD group performs better). The studies were conducted in a number of different countries and in various languages (English, Hebrew, Korean, Mandarin Chinese, Polish, Spanish), and included participants spanning a reasonably broad age range: from mid-childhood through adolescence to young adult samples. Participants' mean IQ scores ranged from average to above average, which can be expected given that all studies matched the ASD participants to a TD group on verbal- or full-scale IQ. In fact, there was only one study that matched the samples on full-scale IQ. The metaphor tasks were very diverse, including conventional and novel metaphors. Metaphor forms also varied and sometimes included sets of examples that border what we could call metaphor proper, such as *adjective + noun* collocations (e.g., Gold et al., 2010; Gold & Faust, 2010; Kasirer & Mashal, 2016) and proverbs (Yi et al. 2013). One study (Landa & Goldberg, 2005) included other types of figurative language besides metaphors. The materials were presented in auditory, written or pictorial formats. The responses that were required also varied, including selecting a picture or a sentence from multiple options, verifying the plausibility or meaningfulness of metaphoric and non-metaphoric expressions, and providing verbal explanations regarding the meaning of metaphoric expressions.

Fig. 2 presents the effect size of the group differences in metaphor processing (Hedges' g with 95% CIs) between individuals with ASD and matched TD controls. Overall, the results showed a medium-to-large group difference in metaphor processing ($g = 0.76$, $p < .001$; 95% CI: 0.50–1.01). Nevertheless, it is notable that in 7 out of the 15 studies included in the analysis, the effect size of group difference was not significantly different from zero. There was only one study that matched the ASD and TD groups on full-scale IQ (Borkowska, 2015). This study reported the largest group difference in metaphor processing. When we excluded this study, the effect size of group difference was $g = 0.67$ ($p < .001$; 95% CI: 0.46–0.89).

The heterogeneity between studies was significant ($Q(14) = 36.17$, $p = .001$, $I^2 = 61.29$). Given that a considerable proportion of the variance in effect sizes appeared to be attributable to moderator variables, we conducted a meta-regression analysis with mean

aracteristics, type of metaphor task and performance on metaphor task in the studies included in the meta-analysis. The studies are listed in increasing order of the effect size of group

details, country and language	N ASD (TD)	Mean age ASD (TD)	Mean VIQ ASD (TD)	Task type
al. (2020); UK (English)	23 (27)	23.13 (22.26)	116 (106.1) ¹	conventional and novel metaphor verification, written presentation
010); Israel (Hebrew)	16 (16)	21.9 (23.1)	105.8 (106.8)	conventional and novel metaphor verification, written presentation
al. (2016) child; Spain (Spanish)	25 (19)	11.3 (11.9)	110.71 (105.76) ²	matching written sentences to one of four possible images, including distractor
al. (2016) adult; Spain (Spanish)	20 (20)	18.1 (18.3)	118.3 (118.34) ²	matching written sentences to one of four possible images, including distractor
shal (2016); Israel (Hebrew)	34 (39)	12.59 (12.26)	matched on verbal ability	conventional and novel metaphor understanding and generation, written presentation
(2010); Israel (Hebrew)	27 (36)	22.95 (24.66)	101.52 (100.17)	conventional and novel metaphor verification, written presentation
3); Korea (Korean)	12 (12)	8.26 (8.19)	97.17 (100.33)	conventional and novel proverb verification (multiple choice), written presentation
2015); China (Mandarin Chinese)	15 (15)	6.48 (6.28)	109.6 (110.87)	conventional and novel metaphors embedded in stories, written presentation
al. (2011); USA (English)	30 (28)	12.33 (12.5)	104.4 (108.5)	metaphor understanding and explanation (selecting correct picture that corresponds to the experimenter's metaphoric description)
lberg (2005); USA (English)	19 (19)	11.01 (11)	113.5 (115.6)	metaphor explanation, verbal presentation
man (2017); Israel (Hebrew)	32 (32)	9.33 (9.33)	102 (102) ³	verbalizing a metaphor presented in pictures
(2015); Chile (Spanish)	5 (5)	26.6 (27.6)	matched on VIQ	novel metaphor interpretation and MCO, verbal presentation
(2002); USA (English)	8 (8)	16.25 (16.88)	111 (116)	metaphor-picture matching and rating novel metaphors as plausible or implausible, written presentation
l. (1995); USA (English)	62 (50)	17.79 (16.91)	94.06 (98.76)	making inferences on the basis of ambiguous metaphoric expressions, verbal presentation
015); Poland (Polish)	23 (28)	10.84 (10.90)	93.10 (95.2) ⁴	verbal and pictorial metaphor verification and explanation

participants were matched on their raw scores of the WAIS-II vocabulary subscale. Full-scale IQs of the participants were estimated on the basis of this single subtest, which can only give a rough

IQ reported, but the groups were also matched on various verbal skills (but VIQ not reported separately).

IQ not reported in paper, but ASD and TD groups were pair-matched on verbal IQ.

IQ reported.

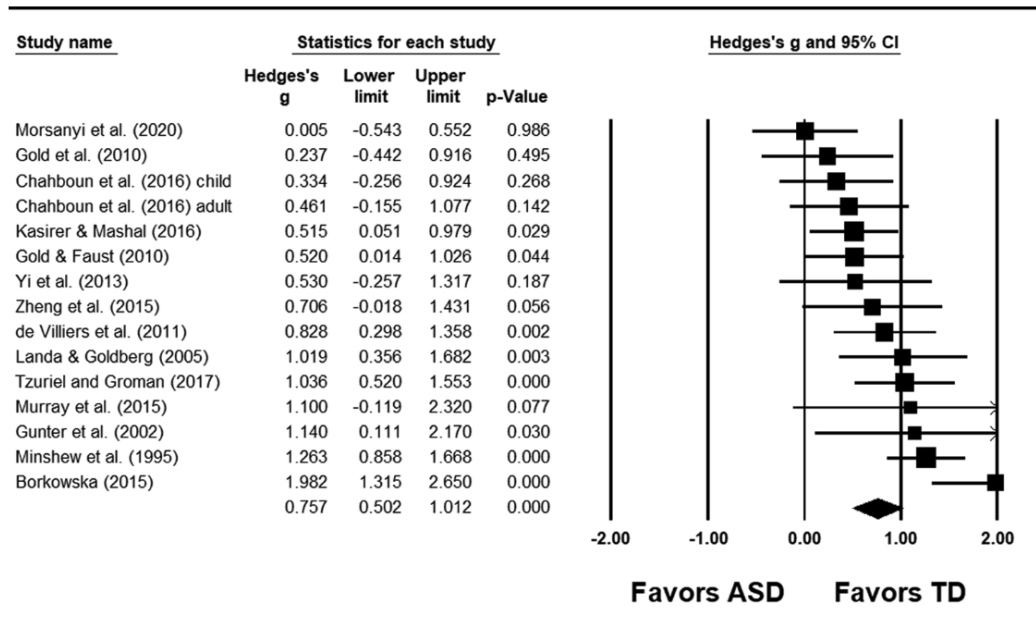


Fig. 2. Hedges' *g* effect sizes with 95% confidence intervals for group differences in metaphor processing accuracy between individuals with ASD and age- and verbal intelligence-matched controls. The overall mean effect size is presented in the bottom line (and marked by \blacklozenge in the figure).

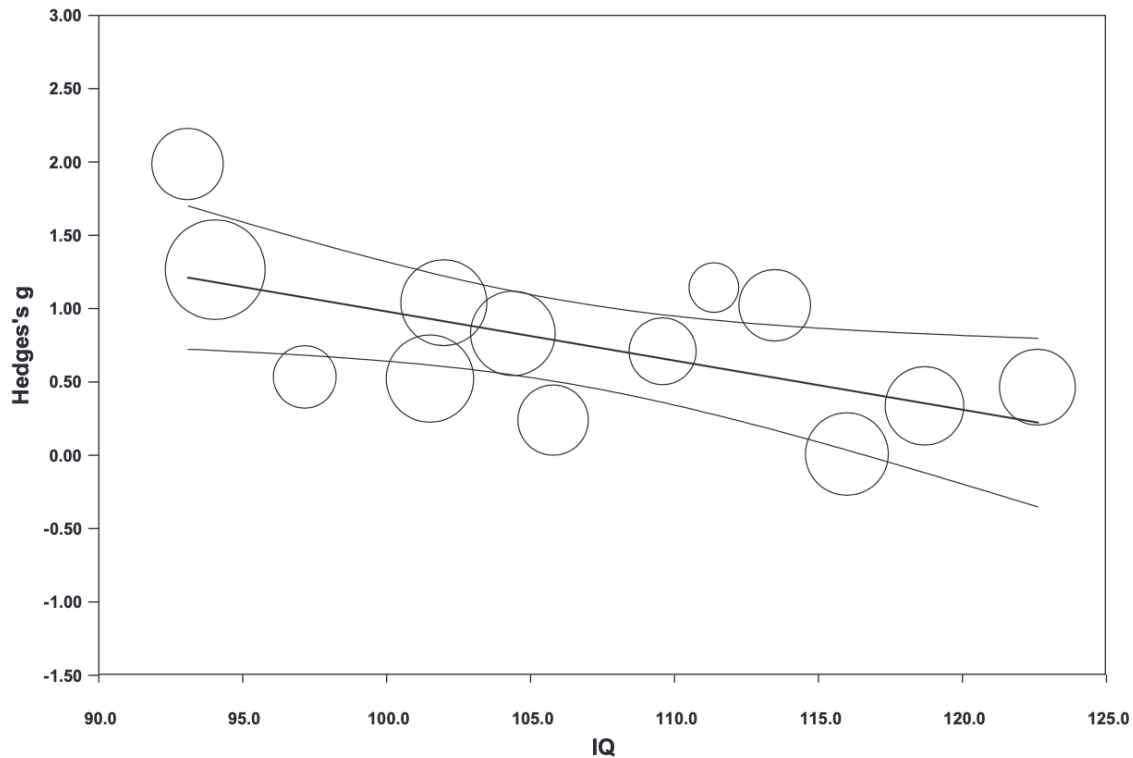


Fig. 3. Plot presenting the regression of Hedges' *g* effect sizes for group differences in metaphor processing scores on the average verbal IQ of the ASD participants in each study.

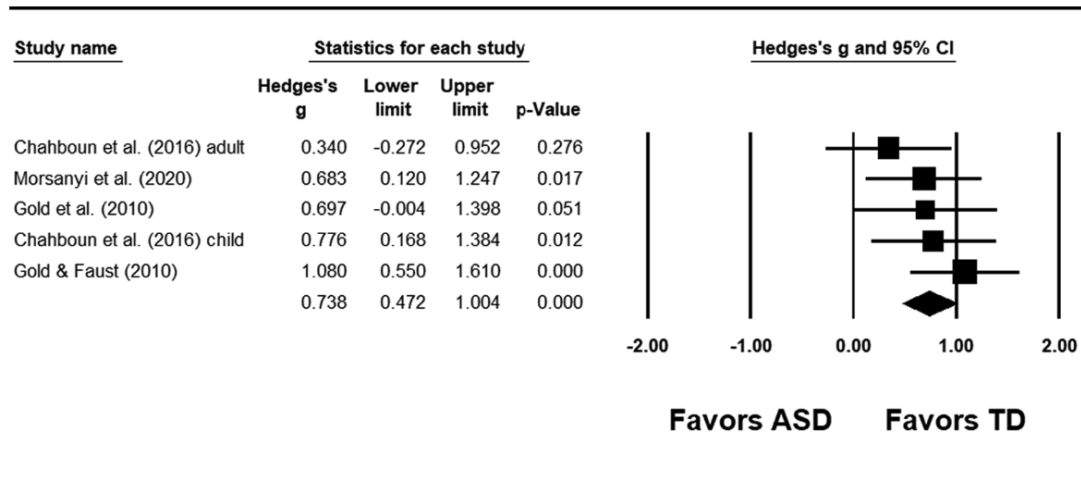


Fig. 4. Hedges' g effect sizes with 95% confidence intervals for group differences in metaphor processing reaction times between individuals with ASD and age- and verbal intelligence-matched controls. The overall mean effect size is presented in the bottom line (and marked by \blacklozenge in the figure).

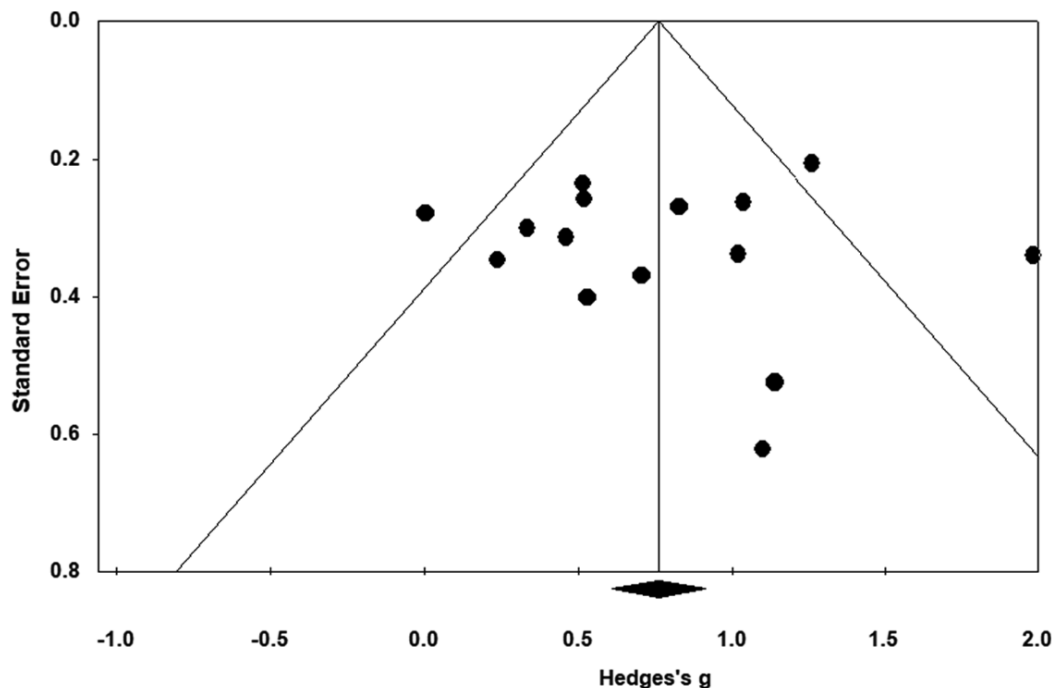


Fig. 5. Funnel plot to assess the potential impact of publication bias.

($Q(2) = 13.92, p < .001$), explaining 80% of the variance in effect sizes between studies. The mean verbal IQ of autistic participants was significantly negatively related to the effect size of group differences ($p = .002$; see Fig. 3).³ There was also a non-significant trend for a negative relationship between participants' age and the effect size of group differences ($p = .068$).^{4 4}

Five studies in our review (Chahboun et al., 2016, child and adult samples; Gold et al., 2010; Gold & Faust, 2010; Morsanyi, Hamilton & Holyoak, 2020) also reported reaction time results related to metaphor processing, besides accuracy. Although reaction times are more difficult to interpret than accuracy data, and this was not our main focus of interest, we also present a meta-analysis of these results. The overall estimate of effect size was remarkably similar to the estimate based on accuracy data ($g = 0.76, p < .001$; 95% CI: 0.47–1.00), indicating a medium-to-large group difference in metaphor processing (Fig. 4).

Fig. 5 presents a funnel plot to assess the possible impact of publication bias. This analysis included all studies ($n = 15$) that met our original inclusion criteria. Our analysis, using the fail-Safe N statistic, revealed that 334 missing studies would be needed to change the overall result of the meta-analysis to not significantly different from zero. This is well above the critical value of 85, which was the threshold for the results to be considered stable based on Rosenthal's (1969) approach. Nevertheless, this method is less reliable when the number of studies included in the analysis is low (Lau et al., 2006). In the funnel plot, the study that matched participants on full-scale IQ and reported a very large effect size for group differences appears as an outlier. In addition, there were two studies with large effect sizes and a small number of participants. As a final check for the possible presence of publication bias, we also computed the correlation between sample sizes and effect sizes. The correlation was positive and non-significant ($r(13) = 0.12$, $p = .669$), indicating that smaller studies were not systematically associated with larger effect sizes.

In terms of the quality of the papers (see Table A1 in the Appendix for further details), we judged that all papers were of at least moderate quality, with most of them judged to fall in the moderate to strong category. Thus, apart from some minor design issues or omitted details, the methodology of the papers appears to be robust. The papers usually matched the groups at least on age and gender, and several papers also matched autistic and non-autistic participants on educational background and socio-economic status. Exclusions and withdrawals were not typically described in the studies, although some papers used some specific inclusion criteria (for example, that the participants in the ASD group had to have an IQ above 70). We list the specific diagnostic categories of the participants in the studies in Table A1 (although in the DSM-5, all of these categories have been merged into a single autism spectrum disorder diagnostic category). Most studies used metaphor tasks that were created by the authors, but these tasks were typically piloted with independent samples, indicating that the tasks can be judged to be appropriate for the relevant age and ability-level groups. Unfortunately, most studies did not report the reliability of the metaphor task, although the number of items tended to be reasonably large. In addition, some studies used existing standardized instruments to assess metaphor processing. All studies matched the participants on a standardized test of verbal ability or IQ (usually using a version of the Wechsler intelligence scale), and some studies assessed language skills with multiple tests. Apart from two studies (one published as a conference poster, and one in preparation for publication), all studies were published in peer-reviewed journals.

Discussion

Overview

The standard diagnostic criteria for ASD include impairments in figurative language understanding. Metaphor is a particularly challenging form of figurative language, because it involves finding connections between distinct semantic domains. Some theorists have argued that although autistic people, as a group, are indeed characterized by impairments in nonliteral language processing, this problem is related to more general issues with language skills in autism. Thus, when autistic and non-autistic people are matched on verbal ability, these group differences should disappear (e.g., Brock et al., 2008; Gernsbacher & Pripas-Kapit, 2012; Geurts et al., 2020; Kalandadze et al., 2018; Norbury, 2005). It has also been suggested that there may be a developmental delay in metaphor processing for autistic people, with group differences decreasing with age, and in some cases disappearing in adulthood (Saban-Bezalel & Mashal, 2018). However, others have maintained that group differences in metaphor processing remain even for high-functioning autistic adults (Vulchanova et al., 2015).

Here we investigated these issues by performing a systematic literature review and meta-analysis. We selected studies in which both chronological age and IQ (ideally, verbal IQ) were equated across an ASD group and a TD control group. Our goal was to determine whether group differences still exist once chronological age and intellectual ability (more specifically, verbal intelligence) is equated across groups. We also conducted meta-regression analyses to investigate the relations between participants' age and IQ and the effect size of group differences.

Overall, we found a medium-to-large group difference in metaphor processing in autism based on accuracy measures (Hedges' $g = 0.76$, indicating greater accuracy for TD than ASD participants; 0.67 if we exclude a single study that matched the groups on full-scale IQ rather than verbal IQ). Interestingly, the present estimated effect size is very similar to that ($g = 0.63$) reported by Kalandadze et al. (2019), even though the selection criteria used by the two studies differed. (Kalandadze et al. did not restrict their analysis to studies that matched the groups on both age and IQ, and there was also little overlap between the studies included in the two meta-analyses.) The present study also found an overall advantage for TD groups using reaction time as the dependent measure (Hedges' $g = 0.74$), based on the results of five studies that reported both accuracy and RTs for metaphor processing. This estimate is remarkably similar to the overall effect size for accuracy. Although the number of studies included in the analysis was modest, it was similar to recent meta-analyses related to autism research (e.g., Griego et al., 2019; Kalandadze et al., 2019). Moreover, the quality of the studies was generally good, and there was no evidence for a significant publication bias, which suggests that our estimate of the overall effect size may be considered reliable.

Across studies included in our meta-analysis, we found reliable heterogeneity in effect sizes for group differences in accuracy. In all studies the mean difference favored the TD group; however, in seven of the individual studies (almost half of the studies included in our meta-analysis) the group difference was not statistically reliable. This heterogeneity suggests that impairments in metaphor processing may vary across types of metaphor processing tasks, or across certain subgroups within the autistic population. One

the seven studies that did not find a reliable group difference used verification tasks (Chahboun et al., 2016, child and adult sample; Gold et al., 2010; Morsanyi et al., 2020), and two used a multiple-choice response format (Yi et al., 2013; Zheng et al., 2015). These paradigms minimize the need to exercise social or verbal skills when formulating a response. The other study that yielded a non-significant group difference (Murray et al., 2015) did require a verbal response. However, the effect size in this study was in fact large, and the lack of significance is likely attributable to the small sample size used (only five participants in each group).

By contrast, the studies that reported larger and more reliable group differences generally used tasks that required more complex responses, such as verbalizing a metaphor presented in pictures (Tzuriel & Groman, 2017), explaining the meaning of metaphors (Borkowska, 2015; de Villiers et al., 2011; Landa & Goldberg, 2005), or making inferences on the basis of ambiguous metaphoric expressions (Minshew et al., 1995). The general tendency for autistic people to find it challenging to provide a verbal explanation of metaphors is consistent with results of the meta-analysis reported by Kalandadze et al. (2019).

Based on previous findings, we examined the role of verbal intelligence and chronological age as potential moderators of the effect size of group differences in accuracy. We found a significant effect of verbal intelligence, such that group differences in metaphor processing were smaller (or non-existent) for participants with higher verbal skills. Indeed, about 60% of the variance in effect sizes between studies was explained by variations in participants' verbal ability. It is also notable that the only study that matched the groups on full-scale IQ (rather than verbal IQ) showed the largest group difference (Borkowska, 2015). Given the typical uneven cognitive profiles of people with ASD (e.g., Stevenson & Gernsbacher, 2013), it is possible that matching on full-scale IQ might mask a disadvantage for the ASD group on the verbal component of IQ tests.

In addition, even when groups are matched on verbal IQ, group differences may exist for some types of verbal skill. Standard measures of verbal IQ focus on recognition of similarities, understanding of concepts, and logical reasoning. As such, they may not adequately measure syntactic ability and other core language skills (except for vocabulary). Although some studies in our analysis matched autistic and non-autistic participants on core language skills, others included global measures of verbal intelligence, without giving details of performance on individual subtests (see details of the tasks used for matching the groups in Table A1 in the Appendix). It is important to distinguish between different structural language skills (e.g., grammar, word knowledge, and semantic operations), because semantic processing may be more directly involved in metaphor processing than syntax. As suggested earlier, ASD may impair the ability to perform the semantic operations involved in conceptual combination, and metaphor may place especially great demands on this mechanism. Another possibility is that verbal reasoning abilities (rather than core language skills) might be more relevant to determining group differences in metaphor processing (although this is not a typical assumption in the literature, and no study in our sample matched their participants specifically on these skills).

It is noteworthy that among the studies included in our meta-analysis, Chahboun et al. (2016) reported that (in their adult sample) the ASD group scored lower on receptive grammar. Landa and Goldberg (2005) found that the ASD group in their study scored lower on the CELF-R formulated sentences subtest. Borkowska (2015) reported that the ASD and TD groups in her study differed on several subscales of the Right Hemisphere Language Battery (the Test of Inferred Meaning Comprehension, Humor Appreciation, Emotional Prosody, and Linguistic Prosody subtests; Bryan, 1989), in addition to the metaphor-related tasks, with the ASD group scoring lower on all of these subtests. Finally, Minshew et al. (1995) found that the groups in their study differed on the Word Attack subtest of the Woodcock Reading Mastery Test—Revised (WRMT-R) (Woodcock, 1987), with the ASD group scoring higher in this case.

These results show that even when the groups are well-matched on overall verbal intelligence, this does not guarantee that they are matched on all important aspects of language. For example, the results of Borkowska (2015) suggest that their ASD group had a general weakness in inferring meaning from ambiguous expressions, rather than a specific problem with understanding metaphorical meaning. Indeed, there is evidence that people with ASD have greater difficulty in interpreting the fixed meanings of idioms (especially those based on cultural knowledge) than in interpreting novel metaphors (Chahboun et al., 2016; Vulchanova, Milburn, Vulchanov, & Baggio, 2019).

Another possible factor that might be related to the size of group differences is the chronological age of participants (Saban-Bezalel & Mashal, 2018). We found a non-significant trend for a negative relationship between chronological age and the effect size of group differences. It is plausible that impairments in figurative language processing decrease with age, as the necessary cognitive and linguistic components develop, along with possible compensatory strategies. However, the available data do not support a strong conclusion about the effect of age.

Relations between metaphor processing, analogical reasoning, and linguistic skills

The findings in the present meta-analysis of studies of metaphor processing in ASD stand in contrast to those that emerged in a similar meta-analysis of studies of analogical reasoning in ASD (Morsanyi, Stamenković, & Holyoak, 2020). The present meta-analysis revealed overall reduced accuracy and slower responding on metaphor tasks for ASD groups relative to TD groups matched in verbal IQ and chronological age. In contrast, a meta-analysis for analogical reasoning found that ASD groups performed as well as TD groups matched in overall IQ and age, with ASD groups showing small but reliable advantages compared to TD groups for formal visual analogies (Ravens Progressive Matrices). Moreover, whereas the current meta-analysis found that autistic participants with lower levels of (verbal) intelligence were particularly likely to show impairments in metaphor processing compared to TD controls matched in verbal IQ, Morsanyi et al. (2020) found that ASD participants with lower general intelligence outperformed IQ-matched TD controls

These qualitative differences in ability patterns between metaphor and analogy tasks for people with ASD cast doubt on theories that assume metaphor processing primarily relies on analogical reasoning (Gentner & Clement, 1988; Tourangeau & Sternberg, 1981). Also, if metaphor processing in autism relied heavily on analogical reasoning (a relative strength in ASD), one might expect metaphor to be easier for ASD populations than other forms of figurative language. However, a recent systematic review of figurative language in autism (Kalandadze et al., 2018) found that metaphor processing was relatively challenging for autistic individuals compared to sarcasm and irony. It therefore appears unlikely that analogical reasoning plays a major role in metaphor processing in autism (although it may provide a potential compensatory strategy).

It is noteworthy that studies of analogical reasoning in ASD have almost universally used nonverbal materials, such as pictures and geometric patterns (Morsanyi et al., 2020), whereas studies of metaphor processing focus on linguistic stimuli. As many theorists have argued, the deficits in metaphor processing for people with autism may be closely linked to more general deficits in linguistic skills. For example, even though autistic individuals are generally unimpaired in spatial skills, they show deficits in the use of spatial language (Bochynska, Coventry, Vulchanov, & Vulchanova, 2020), and the integration of materials presented in different modalities (for example, pictorial and verbal) might pose a particular challenge for them (cf. Kalandadze et al., 2019). The present meta-analysis selected studies in which ASD and TD groups were matched in verbal IQ, and nonetheless found an overall deficit for ASD groups. But as noted above, verbal IQ was a strong moderator of group differences, which were minimal for studies that tested people with high verbal intelligence. Moreover, as Gernsbacher and Pripas-Kapit (2012) have emphasized, standard measures of verbal IQ (which rely heavily on tests that assess vocabulary knowledge and understanding of individual concepts) do not provide a full assessment of linguistic skills. Metaphor processing, especially when the task demands generation of complex verbal responses, likely depends on a wide range of syntactic as well as semantic knowledge, coupled with skill in drawing inferences and performing information integration. The present meta-analysis is thus compatible with evidence that metaphor processing in ASD is intimately linked to general linguistic skills (e.g., Vulchanova et al., 2015).

Limitations and future directions

Although the current meta-analysis helps to resolve some debates in the literature on metaphor processing in autism, it has a number of limitations. The small number of studies that met our inclusion criteria is problematic because we had limited statistical power to evaluate the effects of potential mediator variables. In particular, our findings relating to the effect of chronological age on group differences were inconclusive.

We selected studies that matched groups in verbal IQ; but as noted above, studies that include broader assessments of linguistic ability would be especially valuable. The studies included in our meta-analysis were extremely diverse in the nature of the tasks used to assess metaphor processing. Our findings suggest that people with ASD exhibit greater deficits relative to TD controls for tasks that require more complex verbal comprehension and/or production. To systematically investigate the impact of properties of the task on group differences, it would be useful to run studies in which participants perform multiple tasks involving metaphor processing. It would be highly desirable for researchers to report measures of the reliability of their metaphor tasks, and also to make the full sets of materials available to other researchers (in a similar manner to a recent taxonomy of verbal analogy tests; Ichien, Lu, & Holyoak, 2020). Studies that combine a broad assessment of linguistic skills with metaphor tasks varying in their linguistic complexity (and ideally, also with tasks involving other varieties of nonliteral language) would help to identify key components of verbal skills that are critical in metaphor processing, both in ASD and TD populations.

Another important direction for future research will be to combine metaphor tasks with assessments of other cognitive skills (in addition to linguistic skills). For example, to more directly determine the relation between metaphor processing and analogical reasoning, studies could administer assessments of both abilities to ASD and matched TD control groups (cf. Morsanyi et al., 2020). Similarly, more studies should combine tests of metaphor with assessments of theory of mind and of executive functions. Theoretical progress can also be advanced by studies that relate information processing in autism to other cognitive tasks involving semantic processing, conceptual integration, and the ability to make gist-based inferences (see Miller, Odegard & Allen, 2014 for a detailed discussion). In general, further progress in understanding the nature of metaphor processing in people with ASD will require larger-scale studies that provide broader assessments of individual differences in linguistic and cognitive abilities.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgments

Preparation of this paper was supported by two project grants to Dušan Stamenković (project 360/1-16-1-01 of the Faculty of Philosophy, University of Niš, and project 0-18-18 of the Serbian Academy of Arts and Sciences), and by NSF Grant BCS-1827374 to

Characteristics, design and properties of the metaphor task, materials used for matching the groups on (verbal) IQ and publication status of the studies included in the meta-analysis, with global

Study ID	Sample characteristics	Design and properties of the metaphor task	Materials used for matching the groups on IQ	Publication status	Global quality rating
1	High functioning autism and Asperger's syndrome; both ASD and control group recruited from universities; no withdrawals; groups matched on age, gender, educational background and verbal ability.	Task created by authors, based on Happé, 1995 and Norbury, 2005; 20 metaphors; Cronbach's alpha: 0.72	WASI-II Vocabulary subscale	manuscript in preparation	moderate to strong
10	Asperger's syndrome; ASD recruited from protected homes or via an Asperger's association; one participant with ASD excluded; groups matched on age, gender and VIQ.	Task created by authors; 120 metaphors; piloted with independent samples	Task(s) used for matching on VIQ not described.	journal paper	moderate to strong
11	High-functioning autism (confirmed with ADOS); recruitment details and exclusions or withdrawals not described; groups matched on age, gender and verbal ability.	Task created by authors; 20 metaphors; piloted with independent samples	WISC-IV similarities, vocabulary and verbal comprehension subtests; reading comprehension and receptive grammar (Spanish adaptation of TROG)	journal paper	moderate to strong
12	High-functioning autism (confirmed with ADOS); recruitment details and exclusions or withdrawals not described; groups matched on age, gender and verbal ability.	Task created by authors; 20 metaphors; piloted with independent samples	WAIS-IV similarities, vocabulary and verbal comprehension subtests; reading comprehension and receptive grammar (Spanish adaptation of TROG)	journal paper	moderate to strong
13	Autistic disorder, Asperger's and PDD-NOS; ASD and control children recruited from the same regular elementary schools; participants excluded if they did not score within the age-appropriate range on the screening tests; groups matched on age, gender, verbal ability and educational background.	Task created by authors; 20 metaphors; piloted with independent samples	WISC-IV vocabulary subtest; Hebrew picture naming test (Kavé, 2005)	journal paper	strong
14	Asperger's syndrome; controls recruited from universities; exclusions or withdrawals not described; groups matched on age, VIQ, handedness (marginal difference in gender distribution).	Task created by authors; 60 metaphors; piloted with independent samples	WAIS-III verbal IQ	journal paper	moderate to strong
15	Asperger's syndrome; recruitment process and exclusions or withdrawals not described; groups matched on age, gender, VIQ.	Task created by authors; 20 proverbs; piloted with independent samples	receptive and expressive vocabulary test (Kim, Hong, Kim, Jang, & Lee, 2009), Korean sentence comprehension test (Pac, Lim, Lee, & Jang, 2004), Revised Chinese version of WISC verbal IQ	journal paper	moderate
16	High-functioning autism; children with ASD recruited from special schools; no exclusions or withdrawals; groups matched on age, VIQ, handedness.	Task created by authors; 12 metaphors; piloted with independent samples		journal paper	moderate to strong

(continued on next page)

tinued)

etails	Sample characteristics	Design and properties of the metaphor task	Materials used for matching the groups on IQ	Publication status	Global quality rating
al.	High-functioning autism or PDD-NOS; recruitment process and exclusions or withdrawals not described; groups matched on age and verbal ability.	Task created by the authors, number of items and reliability not reported	VIQ measured by DAS or Wechsler Intelligence test	conference poster	moderate
lberg	High-functioning autism; recruited via autism society and special autism research unit; groups individually matched on age, gender and IQ; exclusions or withdrawals not described	Figurative Language subtest of the Test of Language Competence (TLC; Wiig & Secord, 1989): comprehension and interpretation of metaphoric expressions and figures of speech	WISC-R; WISC-III or WAIS-R	journal paper	moderate to strong
man	High-functioning autism and Asperger's; recruited via an autism society or from mainstream schools; exclusions or withdrawals not described; groups pair-matched on age, gender, VIQ and socioeconomic status.	Children's Proverbial Understanding Test (Tzuriet & Valdman, 2009): Metaphorical Mapping subtest, 11 items. Children's Metaphorical Construction Test (Tzuriet, Yosef, & Valdman, 2008): 13 items.	WISC-VI Vocabulary subtest	journal paper	strong
(2015)	Asperger's syndrome; exclusions or withdrawals not described; groups pair-matched on age, gender, level of education, occupation or subject of study, associated diagnoses, medications, handedness and verbal ability.	Spanish version of the Protocole Montréal d'Évaluation de la Communication (Ferrerres et al., 2007) metaphor subtest	Vocabulary and Ideally Complex Item subtests of the Boston Naming Test (Goodglass & Kaplan, 2005); syntax subtest of the Bateria de Lenguaje Objetiva y Criterial (Puyuelo, Wiig, Renom & Solanas, 1997)	journal paper	moderate to strong
(2002)	Asperger's syndrome; recruited via a child psychiatry unit; exclusions or withdrawals not described; groups matched on age and VIQ	Right Hemisphere Language Battery (Bryan, 1989): metaphor-picture matching, written metaphor choice subtests	WISC-R for the ASD group, British Picture Vocabulary Scale (Dunn, Dunn, Whetton, & Pintilie, 1982) or the National Adult Reading Test (Nelson, 1991) for control group	journal paper	moderate
1.	High-functioning autism; inclusion and exclusion criteria described; the control group consisted of community volunteers; groups matched on age, intellectual ability, gender, race, and socioeconomic status of the family of origin.	Test of Language Competence (TLC; Wiig & Secord, 1985): Understanding Metaphoric Expressions subtest	WAIS-R or WISC-R verbal IQ	journal paper	strong
(2015)	High-functioning autism; recruited from a special school; inclusion criteria described	Right Hemisphere Language Battery (Bryan, 1989): metaphor-picture matching and explaining, written metaphor choice and explanation subtests	Task(s) used for matching on full scale IQ not described	journal paper	moderate

References

(* indicates studies included in the meta-analysis)

- Baron-Cohen, S., Tager-Flusberg, H., & Cohen, D. J. (Eds.). (2000). *Understanding other minds: Perspectives from developmental cognitive neuroscience* (2nd ed.). Oxford: Oxford University Press.
- Bochynska, A., Coventry, K. R., Vulchanov, V., & Vulchanova, M. (2020). Tell me where it is: Selective difficulties in spatial language on the autism spectrum. *Autism*, OnlineFirst, June 4, 2020. DOI: 10.1177/1362361320921040.
- Borenstein, M., Hedges, L. V., Higgins, J., & Rothstein, H. R. (2009). *Introduction to meta-analysis*. Hoboken, NJ: Wiley. DOI:10.1002/9780470743386.
- Borkowska, A. R. (2015). Language and communicative functions as well as verbal fluency in children with high-functioning autism. *Journal of Intellectual Disability – Diagnosis and Treatment*, 3, 147–153. <https://doi.org/10.6000/2292-2598.2015.03.03.4>.
- Bowdle, B., & Gentner, D. (2005). The career of metaphor. *Psychological Review*, 112, 193–216. <https://doi.org/10.1037/0033-295X.112.1.193>.
- Brock, J., Norbury, C., Einav, S., & Nation, K. (2008). Do individuals with autism process words in context? Evidence from language-mediated eye-movements. *Cognition*, 108, 896–904. <https://doi.org/10.1016/j.cognition.2008.06.007>.
- Brosnan, M., Chapman, E., & Ashwin, C. (2014). Adolescents with autism spectrum disorder show a circumspect reasoning bias rather than ‘jumping-to-conclusions’. *Journal of Autism and Developmental Disorders*, 44, 513–520. <https://doi.org/10.1007/s10803-013-1897-5>.
- Bryan, K. L. (1989). Language prosody and the right hemisphere. *Aphasiology*, 3, 285–299.
- Chahboun, S., Vulchanov, V., Saldaña, D., Eshuis, H., & Vulchanova, M. (2016). Can you play with fire and not hurt yourself? A comparative study in figurative language comprehension between individuals with and without autism spectrum disorder. *PLoS ONE*, 11, Article e0168571. <https://doi.org/10.1371/journal.pone.0168571>.
- Conn, V. S., Valentine, J. C., Cooper, H. M., & Rantz, M. J. (2003). Grey literature in meta-analyses. *Nursing Research*, 52, 256–261.
- Cooper, H., Hedges, L. V., & Valentine, J. C. (2009). *The handbook of research synthesis and meta-analysis* (2nd ed.). New York: Sage.
- Dawson, M., Soulières, I., Gernsbacher, M. A., & Mottron, L. (2007). The level and nature of autistic intelligence. *Psychological Science*, 18, 657–662. <https://doi.org/10.1111/j.1467-9280.2007.01954.x>.
- De Martino, B., Harrison, N. A., Knafo, S., Bird, D., & Dolan, R. J. (2008). Explaining enhanced logical consistency during decision making in autism. *Journal of Neuroscience*, 28, 10746–10750. <https://doi.org/10.1523/JNEUROSCI.2895-08.2008>.
- de Villiers, P. A., de Villiers, J. G., Diaz, S., Cheung, C., Alig, R., Raditz, V., & Paul, R. (2011). *Non-literal language and Theory of Mind in Autism Spectrum Disorders*. San Diego, CA: American Speech-Language-Hearing Association.
- Dunn, L. M., Dunn, L. M., Whetton, C., & Pintilie, D. (1982). *The British Picture Vocabulary Scale: Manual for the short and long forms*. NFER-Nelson.
- Ferreira, A., Abusamra, V., Martínez-Cuitiño, M., Ska, B., Côté, H., & Joannette, Y. (2007). *Protocolo para la Evaluación de la Comunicación de Montréal (MEC)*. Buenos Aires: Neuropsi.
- Frith, U. (2003). *Autism: Explaining the enigma*. Hoboken, NJ: Blackwell.
- Gentner, D. (1988). Metaphor as structure mapping: The relational shift. *Child Development*, 59, 47–59. <https://doi.org/10.2307/1130388>.
- Gentner, D., & Clement, C. (1988). Evidence for relational selectivity in the interpretation of analogy and metaphor. In G. H. Bower (Vol. Ed.), *Advances in the psychology of learning and motivation: Vol. 22*, (pp. 307–358). New York: Academic Press.
- Gernsbacher, M. A., & Pripas-Kapit, S. R. (2012). Who's missing the point? A commentary on claims that autistic persons have a specific deficit in figurative language comprehension. *Metaphor and Symbol*, 27, 93–105. <https://doi.org/10.1080/10926488.2012.656255>.
- Geurts, B., Kissine, M., & van Tiel, B. (2020). Pragmatic reasoning in autism. In K. Morsanyi, & R. Byrne (Eds.). *Reasoning and decision making in autism* (pp. 113–134). London: Routledge.
- Glucksberg, S., & Keysar, B. (1990). Understanding metaphorical comparisons: Beyond similarity. *Psychological Review*, 97, 3–18. <https://doi.org/10.1037/0033-295X.97.1.3>.
- Glucksberg, S., Gildea, P., & Bookin, H. B. (1982). On understanding nonliteral speech: Can people ignore metaphors? *Journal of Verbal Learning and Verbal Behavior*, 21, 85–98. [https://doi.org/10.1016/S0022-5371\(82\)90467-4](https://doi.org/10.1016/S0022-5371(82)90467-4).
- Gold, R., & Faust, M. (2010). Right hemisphere dysfunction and metaphor comprehension in young adults with Asperger syndrome. *Journal of Autism and Developmental Disorders*, 40, 800–811. <https://doi.org/10.1007/s10803-009-0930-1>.
- Gold, R., Faust, M., & Goldstein, A. (2010). Semantic integration during metaphor comprehension in Asperger syndrome. *Brain & Language*, 113, 124–134. <https://doi.org/10.1016/j.bandl.2010.03.002>.
- Goodglass, H., & Kaplan, E. (2005). *Evaluación de la afasia y de trastornos relacionados. (Test de Boston)*. Madrid: Médica Panamericana.
- Griego, A. W., Datzman, J. N., Estrada, S. M., & Middlebrook, S. S. (2019). Suggestibility and false memories in relation to intellectual disability and autism spectrum disorder: A meta-analytic review. *Journal of Intellectual Disability Research*, 63(12), 1464–1474. <https://doi.org/10.1111/jir.12668>.
- Gunter, H. L., Ghaziuddin, M., & Ellis, H. D. (2002). Asperger syndrome: Tests of right hemisphere functioning and interhemispheric communication. *Journal of Autism and Developmental Disorders*, 32, 263–281. <https://doi.org/10.1023/A:1016326701439>.
- Happé, F. (1993). Communicative competence and Theory of Mind in autism: A test of relevance theory. *Cognition*, 48, 101–119. [https://doi.org/10.1016/0010-0277\(93\)90026-r](https://doi.org/10.1016/0010-0277(93)90026-r).
- Happé, F. (1995). Understanding minds and metaphors: Insights from the study of figurative language in autism. *Metaphor and Symbol*, 10, 275–295. https://doi.org/10.1207/s15327868ms1004_3.
- Happé, F., & Frith, U. (2006). The weak coherence account: Detail-focused cognitive style in autism spectrum disorders. *Journal of Autism and Developmental Disorders*, 35, 5–25. <https://doi.org/10.1007/s10803-005-0039-0>.
- Hedges, L. V. (1981). Distribution theory for Glass's estimator of effect size and related estimators. *Journal of Educational Statistics*, 6, 107–128. <https://doi.org/10.2307/1164588>.
- Hedges, L. V., & Olkin, I. (1985). *Statistical methods for meta-analysis*. Orlando: Academic Press.
- Higgins, J. P. T., & Thompson, S. G. (2002). Quantifying heterogeneity in a meta-analysis. *Statistics in Medicine*, 21, 1539–1558. <https://doi.org/10.1002/sim.1186>.
- 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>.
- Highnam, C., Wegmann, J., & Woods, J. (1999). Visual and verbal metaphors among children with typical language and language disorders. *Journal of Communication Disorders*, 32, 25–35. [https://doi.org/10.1016/S0021-9924\(98\)00027-6](https://doi.org/10.1016/S0021-9924(98)00027-6).
- Holyoak, K. J. (2012). Analogy and relational reasoning. In K. J. Holyoak, & R. G. Morrison (Eds.). *The Oxford handbook of thinking and reasoning* (pp. 234–259). New York: Oxford University Press.
- Holyoak, K. J. (2019). *The spider's thread: Metaphor in mind, brain, and poetry*. Cambridge, MA: MIT Press. DOI: 10.7551/mitpress/11119.001.0001.
- Holyoak, K. J., & Stamenković, D. (2018). Metaphor comprehension: A critical review of theories and evidence. *Psychological Bulletin*, 144, 641–671. <https://doi.org/10.1037/bul0000145>.
- Ichien, N., Lu, H., & Holyoak, K. J. (2020). Verbal analogy problem sets: An inventory of testing materials. *Behavior Research Methods*. <https://link.springer.com/article/10.3758/s13428-019-01312-3>.
- Kalandadze, T., Bambina, V., & Næss, K.-A.-B. (2019). A systematic review and meta-analysis of studies on metaphor comprehension in individuals with autism spectrum disorder: Do task properties matter? *Applied Psycholinguistics*, 40, 1421–1454. <https://doi.org/10.1017/S0142716419000328>.

- Kavé, G. (2005). Standardization and norms for a Hebrew naming test. *Brain and Language*, 92, 204–211.
- Kim, Y., Hong, K. H., Kim, K. H., Chang, H. S., & Lee, J. Y. (2009). *Receptive and Expressive Vocabulary Test (REVT)*. Seoul: Seoul Community Rehabilitation Center.
- Kintsch, W. (2000). Metaphor comprehension: A computational theory. *Psychonomic Bulletin & Review*, 7, 257–266. <https://doi.org/10.3758/BF03212981>.
- Lakoff, G., & Johnson, M. (1980). *Metaphors we live by*. Chicago: University of Chicago Press.
- Landa, R. J., & Goldberg, M. C. (2005). Language, social, and executive functions in high functioning autism: A continuum of performance. *Journal of Autism and Developmental Disorders*, 35, 557–573. <https://doi.org/10.1007/s10803-005-0001-1>.
- Lau, J., Ioannidis, J. P., Terrin, N., Schmid, C. H., & Olkin, I. (2006). The case of the misleading funnel plot. *British Medical Journal*, 333, 597–600. <https://doi.org/10.1136/bmj.333.7568.597>.
- Lecce, S., Ronchi, L., Del Sette, P., Bischetti, L., & Bambini, V. (2019). Interpreting physical and mental metaphors: Is Theory of Mind associated with pragmatics in middle childhood? *Journal of Child Language*, 46, 393–407. [0.1017/S030500091800048X](https://doi.org/10.1017/S030500091800048X).
- Levine, T. R., Asada, K. J., & Carpenter, C. (2009). Sample sizes and effect sizes are negatively correlated in meta-analyses: Evidence and implications of a publication bias against nonsignificant findings. *Communication Monographs*, 76, 286–302. DOI: 10.1080/03637750903074685.
- Long, J. (2001). An introduction to and generalization of the “fail-safe N”. Paper presented at the Annual Meeting of the Southwest Educational Research Association (New Orleans, LA, February 1–3, 2001).
- McAuley, L., Tugwell, P., & Moher, D. (2000). Does the inclusion of grey literature influence estimates of intervention effectiveness reported in meta-analyses? *The Lancet*, 356, 1228–1231. [https://doi.org/10.1016/S0140-6736\(00\)02786-0](https://doi.org/10.1016/S0140-6736(00)02786-0).
- Miller, H. L., Odegard, T. N., & Allen, G. (2014). Evaluating information processing in autism spectrum disorder: The case for fuzzy trace theory. *Developmental Review*, 34, 44–76. <https://doi.org/10.1016/j.dr.2013.12.002>.
- Minshew, N. J., Goldstein, G., & Siegel, D. J. (1995). Speech and language in high-functioning autistic individuals. *Neuropsychology*, 9, 255–261. <https://doi.org/10.1037/0894-4105.9.2.255>.
- Moher, D., Liberati, A., Tetzlaff, J., & Altman, D. G., The PRISMA Group (2009). Preferred reporting items for systematic reviews and meta-analyses: The PRISMA statement. *PLoS Medicine*, 6, e1000097. DOI:10.1371/journal.pmed1000097.
- Morsanyi, K., & Holyoak, K. J. (2010). Analogical reasoning ability in autistic and typically-developing children. *Developmental Science*, 13, 578–587. <https://doi.org/10.1111/j.1467-7687.2009.00915.x>.
- Morsanyi, K., Handley, S. J., & Evans, J. St. B. T. (2010). Decontextualised minds: Adolescents with autism are less susceptible to the conjunction fallacy than typically developing adolescents. *Journal of Autism and Developmental Disorders*, 40, 1378–1388. DOI: 10.1007/s10803-010-0993-z.
- Morsanyi, K., Stamenković, D., & Holyoak, K. J. (2020). Analogical reasoning in autism: A systematic review and meta-analysis. In K. Morsanyi, & R. Byrne (Eds.). *Thinking, reasoning, and decision making in autism* (pp. 59–87). London: Routledge. <https://doi.org/10.4324/9781351060912>.
- Morsanyi, K., Hamilton, J., & Holyoak, K. J. (2020). Analogical reasoning and metaphor processing in autism. Manuscript in preparation.
- Murray, C., Tobar, A., Villablanca, F., & Soto, G. (2015). El componente pragmático en adultos con síndrome de Asperger: Actos de habla indirectos, metáforas y coerción aspectual [The pragmatic component in adults with Asperger syndrome: Indirect speech acts, metaphors, and aspectual coercion]. *Revista de Lingüística Teórica y Aplicada*, 53, 35–58. <https://doi.org/10.4067/S0718-48832015000100003>.
- Nelson, H. E. (1991). *National Adult Reading Test (NART): Test manual (rev.ed.)*. Windsor, UK: NFER-Nelson.
- Nippold, M. A. (2016). *Later language development: School-age children, adolescents, and young adults* (4th ed.). Austin, TX: PRO-ED.
- Norbury, C. F. (2005). The relationship between theory of mind and metaphor: Evidence from children with language impairment and autistic spectrum disorder. *British Journal of Developmental Psychology*, 23, 383–399. <https://doi.org/10.1348/026151005X26732>.
- Pae, S. Y., Lim, S. S., Lee, J. H., & Jang, H. S. (2004). *Korean sentence comprehension test*. Seoul: Seoul Community Rehabilitation Center.
- Pellicano, E., Maybery, M., Durkin, K., & Maley, A. (2006). Multiple cognitive capabilities/deficits in children with an autism spectrum disorder: ‘Weak’ central coherence and its relationship to theory of mind and executive control. *Development and Psychopathology*, 18, 77–98. <https://doi.org/10.1017/S095457940600056>.
- Polanin, J. R., Tanner-Smith, E. E., & Hennessy, E. A. (2016). Estimating the difference between published and unpublished effect sizes: A meta-review. *Review of Educational Research*, 86, 207–236. <https://doi.org/10.3102/0034654315582067>.
- Pouscoulous, N. (2011). Metaphor: For adults only? *Belgian Journal of Linguistics*, 25, 51–79. <https://doi.org/10.1075/bjl.25.04pou>.
- Pouscoulous, N. (2014). “The elevator’s buttocks”: Metaphorical abilities in children. In D. Matthews (Ed.). *Pragmatic development in first language acquisition* (pp. 239–260). Amsterdam: John Benjamins Publishing Company. <https://doi.org/10.1075/tilar.10.14pou>.
- Puyuelo, M., Wiig, E., Renom, J., & Solanas, A. (1997). *Bateria de Lenguaje Objetiva y Criterial (BLOC)*. Barcelona: Masson.
- Rajendran, G., & Mitchell, P. (2007). Cognitive theories of autism. *Developmental Review*, 27, 224–260. <https://doi.org/10.1016/j.dr.2007.02.001>.
- Recanatì, F. (2004). *Literal meaning*. Cambridge, UK: Cambridge University Press.
- Rosenthal, R. (1969). Task variations in studies of experimenter expectancy effects. *Perceptual and Motor Skills*, 29, 9–10. <https://doi.org/10.2466/pms.1969.29.1.9>.
- Rosenthal, R. (1979). The “file drawer problem” and tolerance for null results. *Psychological Bulletin*, 85, 638–641. <https://doi.org/10.1037/0033-2909.86.3.638>.
- Rundblad, G., & Annaz, D. (2010). The atypical development of metaphor and metonymy comprehension in children with autism. *Autism*, 14, 29–46. <https://doi.org/10.1177/1362361309340667>.
- Saban-Bezalel, R., & Mashal, N. (2018). Figurative language comprehension and laterality in autism spectrum disorder. In R. E. Jung, & O. Vartanian (Eds.). *Cambridge handbook of the neuroscience of creativity* (pp. 281–296). Cambridge, UK: Cambridge University Press.
- Scott, F. J., & Baron-Cohen, S. (1996). Logical, analogical and psychological reasoning in autism: A test of the Cosmides theory. *Development and Psychopathology*, 8, 235–245. <https://doi.org/10.1017/S0954579400007069>.
- Searle, J. R. (1983). Metaphor. In A. Ortony (Ed.). *Metaphor and thought* (pp. 83–111). (2nd ed.). Cambridge, UK: Cambridge University Press.
- Spain, D., Sin, J., Linder, K. B., McMahon, J., & Happé, F. (2018). Social anxiety in autism spectrum disorder: A systematic review. *Research in Autism Spectrum Disorders*, 52, 51–68. <https://doi.org/10.1016/j.rasd.2018.04.007>.
- Sperber, D., & Wilson, D. (2012). A deflationary account of metaphors. In D. Wilson & D. Sperber, D. (Eds.), *Meaning and relevance* (pp. 97–122). Cambridge, UK: Cambridge University Press. DOI: 10.1017/CBO9781139028370.007.
- Stevenson, J. L., & Gernsbacher, M. A. (2013). Abstract spatial reasoning as an autistic strength. *PLoS ONE*, 8, Article e59329. <https://doi.org/10.1371/journal.pone.0059329>.
- Thomas, B., Ciliska, D., Dobbins, M., & Micucci, S. (2004). A Process for systematically reviewing the literature: Providing the research evidence for public health nursing interventions. *Worldviews on Evidence-Based Nursing*, 1, 176–184. <https://doi.org/10.1111/j.1524-475X.2004.04006.x>.
- Tourangeau, R., & Sternberg, R. J. (1981). Aptness in metaphor. *Cognitive Psychology*, 13, 27–55. [https://doi.org/10.1016/0010-0285\(81\)90003-7](https://doi.org/10.1016/0010-0285(81)90003-7).
- Tzuriel, D., & Groman, T. (2017). Dynamic assessment of figurative language of children in the autistic spectrum: The relation to some cognitive and language aspects. *Journal of Cognitive Education and Psychology*, 16, 38–63. <https://doi.org/10.1891/1945-8959.16.1.38>.
- Tzuriel, D., & Valdman, S., (2009). The development of proverbial understanding as a function of analogical, metaphoric, and verbal abilities. Paper presented at the 12th International Conference of the International Association for Cognitive Education and Psychology, Osnabruck, Germany.
- Tzuriel, D., Yosef, L., & Valdman, S., (2008). Young Children’s Metaphoric Ability (YCMA) test. Unpublished manuscript.
- Vosniadou, S. (1987). Children and metaphors. *Child Development*, 58, 870–885. <https://doi.org/10.2307/1130223>.
- Vosniadou, S., Ortony, A., Reynolds, R. E., & Wilson, P. T. (1984). Sources of difficulty in the young child’s understanding of metaphorical language. *Child Development*, 55, 1588–1606. [0.2307/1130028](https://doi.org/10.2307/1130028).
- Vulchanova, M., Milburn, E., Vulchanov, V., & Baggio, G. (2019). Boon or burden? The role of compositional meaning in figurative language processing and acquisition. *Journal of Logic, Language and Information*, 28, 359–387. <https://doi.org/10.1007/s10849-019-09282-7>.
- Vulchanova, M., Saldaña, D., Chauboun, S., & Vulchanov, V. (2015). Figurative language processing in atypical populations: The ASD perspective. *Frontiers in Human Neuroscience*, 9, 24. <https://doi.org/10.3389/fnhum.2015.00024>.

K. Morsanyi, et al.

Developmental Review 57 (2020) 100925

- Wisniewski, E. J. (1997). When concepts combine. *Psychonomic Bulletin & Review*, 4, 167–183. <https://doi.org/10.3758/BF03209392>.
- Wojcik, D. Z., Díez, E., Alonso, M. A., Martín-Cilleros, M. V., Guisuraga-Fernández, Z., Fernández, M., ... Fernandez, A. (2018). Diminished false memory in adults with autism spectrum disorder: Evidence of identify-to-reject mechanism impairment. *Research in Autism Spectrum Disorders*, 45, 51–57. <https://doi.org/10.1016/j.rasd.2017.11.001>.
- Woodcock, R. W. (1987). *Woodcock reading mastery test-revised*. MN: AGS; Circle Pines.
- Yi, D., Hwang, M., & Lim, J. A. (2013). Proverb comprehension in children with Asperger's disorder: The role of transparency. *Communication Sciences & Disorders*, 18, 288–296. <https://doi.org/10.12963/csd.13034>.
- Zheng, Q., Jia, Z., & Liang, D. (2015). Metaphor and metonymy comprehension in Chinese-speaking children with high-functioning autism. *Research in Autism Spectrum Disorders*, 10, 51–58. <https://doi.org/10.1016/j.rasd.2014.11.007>.

