



REVIEW

Idiom and proverb processing in autism: a systematic review and meta-analysis

Kinga Morsanyi · Dušan Stamenković

Received: 29 March 2021 / Revised: 10 May 2021 / Accepted: 17 May 2021 / Published online: 29 May 2021
© The Author(s) 2021

Abstract Figurative language processing impairments in autism have been widely demonstrated, and have been considered a defining feature of autism. Studies in this area often consider different types of figurative language together, and less attention has been paid to identifying the factors that might contribute to difficulties in processing specific types of figurative language. Here we present a preregistered systematic review and meta-analysis of studies assessing the comprehension of idioms and proverbs in autism spectrum disorder (ASD) as compared to typically developing (TD) individuals. Idioms are widely used multi-word figurative expressions, which are understood by using contextual information. Proverbs are a related type of fixed, figurative, formulaic expression in a sentential form, typically linked with wisdom. Idioms and proverbs represent forms of figurative language which are more conventionalized and frequently opaquer than metaphors, pointing to a unique way that they are processed in conversational contexts. Our analysis encompassed a total of 11 studies from 10 papers (involving 235

autistic and 224 TD individuals), which met our inclusion criteria (the ASD and TD groups were matched on both chronological age and intelligence). The analysis of accuracy data revealed a group difference favouring the TD over the ASD group, with a medium effect size, and no indication of a publication bias. Participants' age was unrelated to the magnitude of group differences, but there was a trend for smaller group differences in the case of participants with higher (verbal) intelligence. We discuss these findings with reference to theories related to the nature of figurative language impairments in autism.

Keywords Autism spectrum disorder · Idioms · Figurative language · Meta-analysis · Proverbs · Verbal abilities

Introduction

Autism is a neurodevelopmental disorder, which, according to some current estimates (Maenner et al. 2020), affects about 1 in 54 people. Autism is diagnosed on the basis of clinical observations of certain behavioural features, including persistent difficulties in social communication and interaction, and repetitive and stereotyped behaviour, interests and activities, which cause clinically significant problems in social, educational or occupational functioning

K. Morsanyi (✉)
Centre for Mathematical Cognition, Schofield Building,
Loughborough University, Ashby Road,
Loughborough LE11 3TT, UK
e-mail: K.E.Morsanyi@lboro.ac.uk

D. Stamenković
Language Cognition Laboratory, Faculty of Philosophy,
University of Niš, Niš, Serbia

(DSM-5, APA, 2013). There are also some characteristic features of autistic cognition, which involve orientation toward detailed perceptual features (often at the expense of central coherence), impairments in social imagination (for example, in imaginative play) and, most importantly for the present discussion, deficits in communicative behaviours (e.g., Bal et al. 2019; Happé & Frith, 2006; Lobban-Shymko et al. 2017; Pellicano et al. 2006; Rajendran & Mitchell, 2007).

Nevertheless, starting from the early descriptions of autism (Kanner, 1943), some special skills and talents have also been identified, including performance on visuospatial tasks and visual search (e.g., Dawson et al. 2007; Frith, 1989; Wing & Wing, 1976). In terms of performance on verbal tasks, outstanding performance on some complex tasks relating to reasoning and judgment have also been described. In particular, autistic people have been found to be less susceptible to some reasoning biases and memory illusions than typically developing people (e.g., De Martino et al. 2008; Morsanyi et al. 2010; Wojcik et al. 2018), although these findings may be attributed to impairments in the automatic processing of linguistic information in context (Pijnacker et al. 2010).

Nevertheless, not all forms of contextual processing are impaired in autism. Nonverbal analogical reasoning with various materials (e.g., analogies based on perceptual relations, and scene analogies) has been identified as a particular strength. Specifically, autistic people perform at the same level as age- and IQ-matched typically developing individuals on nonverbal analogy tasks, and autistic people with learning difficulties show superior performance when compared to age- and IQ-matched controls (see Morsanyi et al. 2020a for a systematic review and meta-analysis). Analogical reasoning relies on the ability to find and exploit similarities among entities based on relations, rather than the features of entities (e.g., Gentner, 2010; Holyoak, 2012; Holyoak & Lu, 2021). In other words, it requires flexibility and abstraction, instead of focusing on specific details. Surprisingly, autistic people have been found to be able to perform analogical reasoning even in the presence of salient distractors, and the strategies used to solve analogical reasoning problems have also appeared very similar to the strategies used by typically developing individuals (Morsanyi & Holyoak, 2010).

Impairments in figurative language understanding have long been considered a defining characteristic of autism (e.g., Happé, 1993; 1995). Nevertheless, language impairments are no longer part of the diagnostic criteria of autism (APA 2013), and there is an ongoing debate about the extent to which impairments in figurative language are present in autism, independent of core language impairments (see Kalandadze et al. 2019; Kalandadze et al. 2018; Morsanyi et al. 2020b; Saban-Bezael & Mashal, 2018; Vulchanova et al. 2015 for reviews). Additionally, although in the literature on figurative language in autism it is common practice to investigate various forms of figurative language together, the processing requirements of different figures of speech are not the same. For example, the appreciation of irony and sarcasm necessitates detecting the incongruity between the state of affairs and the literal interpretation of a verbal message. By contrast, comprehending metaphors has been proposed to build on the ability to find and exploit similarities based on relations among entities (e.g., Tourangeau & Sternberg, 1981) as well as the activation and integration of semantic and conceptual information (Kintsch, 2000). In a recent meta-analysis of figurative language processing in autism, Kalandadze et al. (2018) reported that the effect size of group differences in irony and sarcasm (Hedges' $g = 0.48$) appeared smaller than the effect size of group differences in metaphor processing (Hedges' $g = 0.72$). This finding highlights the practical significance of differences in the underlying cognitive mechanisms between different figures of speech.

In light of the above considerations, we present a systematic review and meta-analysis, focusing on idioms and proverbs, both of which represent conventionalized, and often opaque forms of figurative language, which points to a unique way they are processed in conversational contexts. These special types of symbols with (frequently) abstract meanings are likely to depend more intensely on language abilities and crystallized intelligence and less on fluid intelligence, as compared to other forms of figurative language, such as metaphors. Based on these considerations, we have decided to investigate idiom and proverb processing together in the present study.

Idiom and proverb processing

Idioms are multi-word expressions with a non-literal meaning, which can only be derived by using contextual information. Idiomatic expressions are one of the most commonly used forms of figurative language in everyday conversational situations. Just like phrasal verbs, they can be a string of words that corresponds to a single semantic unit (Saeed 2016, p. 56, 444). In most cases, one needs to go beyond their literal meaning in order to understand the essence of what is said (Titone & Connine, 1999). Proverbs are fixed, figurative and traditional expressions, presented in a sentential form and in succinct and formulaic language (Mieder, 2004, p. xi). They typically involve social or moral norms and expressions of wisdom with regard to different everyday life situations. The origins of both of these language phenomena are diverse and can involve metaphoricity, as well as properties of other figures of speech, such as metonymy and hyperbole.

The debate surrounding the processing of idioms is far more complex and longer-running as compared to proverbs. There are two main positions (see Vulchanova et al. 2015) with each of them implying a different approach to the nature of idiom processing, placing them at a different position in terms of processing complexity (see Titone & Connine, 1999; Vulchanova et al. 2011). According to one group of scholars, mostly linked with the lexical representation hypothesis, idioms are stored as lexical items, but along with (usually fast) retrieval, there is also a more complex literal compositional computation process involved in idiom comprehension, as each element is decomposed separately (e.g., Bobrow & Bell, 1973; Chomsky, 1980; Hamblin & Gibbs, 1999; Swinney & Cutler, 1979). The opposing view is based on the configuration hypothesis (Cacciari & Glucksberg, 1991), a more compositional approach, which sees idiom comprehension as a dynamic process in which idioms are considered complex expressions, with constituent parts contributing to the overall meaning of the expression (e.g. Cacciari & Glucksberg, 1991; Cacciari & Tabossi, 1988). In reviewing different positions, Titone and Connine (1999) proposed a middle-ground or hybrid model which combines the compositional and the non-compositional aspects of idiom comprehension in which idioms retain their lexical status along with a degree of decomposability.

Proverbs are similar to idioms in terms of their origins and degree of conventionality, but are more tied to the notion of causality, where language users are expected to be able to use them as “instructive expressions” (Chahboun et al. 2016) in specific contexts. Therefore, they are likely to depend on both pragmatic and analogical abilities, as we apply formulaic language elements to novel situations which emerge in everyday contexts. On the whole, in spite of the fact that we are talking about linguistic elements which are to some degree fossilized (both idioms and proverbs), the use of these elements goes beyond merely understanding their meaning—in many cases, one needs to be able to apply them to concrete contexts, which is why they are tightly bound to pragmatics. As Gibbs and Beitel (1995, p. 133) put it “the ability to correctly explain what a proverb means does not necessarily imply that an individual can think abstractly”. This is in line with the findings that their comprehension is facilitated by the presence of supportive context (see Vulchanova et al. 2015), in which regard they are similar to metaphors (see Gildea & Glucksberg, 1983; Ortony et al. 1978; Stamenković et al. 2020). Nevertheless, idioms (which are sometimes referred to as “dead metaphors”) and proverbs differ from metaphors in that the link between their literal and figurative meaning is indirect, opaque or non-existent (cf., Vulchanova et al. 2019).

Figurative language processing in ASD

Problems related to figurative language comprehension and production have long been seen as one of the characteristic features of autism, and it has been proposed that they might be linked to more general problems with reading other people’s mind (e.g., Baron-Cohen et al. 2000; Happé, 1993, 1995). Indeed, when processing figurative language, the listener has to go beyond the literal meaning of an expression to derive the speaker’s intended meaning. Nevertheless, more recently it has been suggested that a likely cause of figurative language impairments and pragmatic issues in people with autism is a more general language impairment tied with problems with structural language skills and semantic knowledge (e.g., Gernsbacher & Pripas-Kapit, 2012; Geurts et al. 2020; Norbury, 2005; Saban-Bezalel & Mashal, 2018). In support of this claim, in their meta-analysis of several types of figurative language, including idioms and

proverbs, Kalandadze et al. (2018) concluded that individuals with autism exhibited moderately poorer figurative language comprehension skills in comparison to typically developing controls. Nevertheless, in studies in which the ASD and TD groups were matched on verbal ability, differences in figurative language were not found. Kalandadze et al. (2018) did not analyze the results related to idioms and proverbs separately. Indeed, the results extracted from each study often combined performance related to various forms of figurative language, which makes it difficult to assess to what extent this claim applies to these forms of figurative language.

Morsanyi et al. (2020b) performed a systematic review and meta-analysis related to metaphor processing in autism, and also investigated the relations between the age and verbal ability of participants and the effect size of group differences. Unlike the review by Kalandadze et al. (2018), this study only included studies where the autistic and non-autistic groups were matched on both chronological age and (verbal) intelligence. Morsanyi et al. (2020b) found that, overall, there was a difference in metaphor processing ability between autistic and typically developing participants with a medium effect size, even when the groups were matched on age and verbal ability. Nevertheless, group differences were smaller or non-existent in the case of participants with high levels of intelligence. This finding potentially suggests that matching participants on verbal ability might only eliminate group differences in figurative language processing when the groups are characterized by high levels of intellectual functioning. Nevertheless, this meta-analysis was specific to metaphor processing and the findings might not apply to other figures of speech.

Regarding the cognitive processes underlying figurative language processing in autism, in a critical review of the literature, Vulchanova et al. (2015) explains that besides various types of linguistic abilities (ranging from vocabulary size to syntactic and semantic knowledge), individuals' knowledge base, and their ability to draw inferences, perform information integration, suppress irrelevant information, and their mentalizing skills all play a role. In the case of idioms, the role of inferencing skills, the ability to integrate contextual information from both verbal and nonverbal sources, and linguistic skills and competences (especially, the role of syntax—see Whyte et al. 2014) have been highlighted by

Vulchanova et al. (2015). Some properties of idiomatic expressions also affect their processing demands. These include their transparency/decomposability (i.e., transparent expressions are easier to understand than opaque ones), familiarity (i.e., a higher degree of familiarity facilitates understanding), and the context in which they are encountered (cf., Vulchanova et al. 2015). In terms of presentation format, Mashal and Kasirer (2012) found that children with ASD understood more easily visually presented idioms than verbally presented ones, and Vulchanova et al. (2019) reported that idioms and proverbs were processed more easily in a written format as compared to auditory presentation.

In summary, there is good evidence for figurative language impairments in autism, including idioms and proverbs. Nevertheless, this deficit might not always be present. In particular, in the case of high ability autistic participants who are carefully matched to a control group on age and verbal ability, group differences might be absent (see Morsanyi et al. 2020b). Although some reviews of the literature have discussed findings related to idiom and proverb processing in autism (Saban-Bezalel & Mashal, 2018; Vulchanova et al. 2015), no meta-analyses so far have focused specifically on idioms and proverbs. Thus, it is not known whether idiom and proverb processing deficits in autism exist beyond impairments in core language abilities (see Kalandadze et al. 2018). It is also possible that there is a relationship between participants' age and the magnitude of group differences. In particular, if there is a developmental delay in idiom and proverb processing in autism, the effect size of group differences might decrease with age.

Aims and Scope of the Present Review

The present review focuses on idiom and proverb processing in autism. Similar to Morsanyi et al. (2020a, 2020b), this meta-analysis included studies in which the ASD and TD groups were matched on both chronological age and IQ (ideally, verbal IQ). Given that autism is a neurodevelopmental disorder, it is important to consider not only the verbal abilities of participants, but also their age (i.e., the samples should ideally be matched on both age and verbal ability levels). For example, it is not always reasonable to assume that a group of autistic participants with

intellectual impairment would perform on a particular cognitive task in the same way as a typically developing group of younger participants, who are matched on verbal ability to the autistic group (see Jarrold & Brock, 2004 for a detailed discussion of issues around comparing autistic and non-autistic participants on cognitive tasks). Although using this strict criterion eliminates a number of studies from our analyses, this means that we can better understand the potential role of participants' ability levels and their age in the size of group differences. In other words, we can examine the questions of whether there is a developmental delay in idiom and proverb processing in autism, and whether (similar to metaphor processing; see Morsanyi et al. 2020b), group differences disappear in the case of participants with high verbal ability.

Thus, in addition to performing a meta-analysis, we also conducted meta-regressions to determine whether the size of group differences in idiom and proverb processing was related to participants' level of (verbal) intelligence and their chronological age. Additionally, we also performed analyses related to the presence of a publication bias, and evaluated the quality of the studies included in our analyses. Publication bias refers to a tendency to publish studies that report positive (i.e., significant) findings, which can distort the results of meta-analyses, and lead to an overestimation of effect sizes. For example, Polanin et al. (2016), based on a review of meta-analyses published in top-tier educational and psychology journals, found that published studies reported larger effect sizes than unpublished studies, with an average difference of 0.18 standard deviation. Our systematic review and meta-analysis was preregistered in the International Register of Systematic Reviews, PROSPERO, Registration number: CRD42021235762 (available at: https://www.crd.york.ac.uk/PROSPERO/display_record.php?RecordID=235762).

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. The following

databases were included in our search procedure: Web of Science, Scopus, and ProQuest Dissertations & Theses Global for all studies published up to February 2021. We used the following combinations as search terms: *autis* OR ASD OR Asperger** crossed with *idiom* OR proverb**. The target fields in the searches included titles, abstracts, keywords, topics, subjects and indexing. There were no restrictions in terms of the publication year.

The grey literature was also covered, including studies that were not published in peer-reviewed journals (e.g., conference proceedings) and dissertations, which we found in ProQuest Dissertations & Theses Global. By searching for studies in the grey literature, our aim was to minimize the effects of publication bias. We also searched the references of all articles that reached the full-text stage in our search for potential additional papers to be checked against our eligibility criteria.

Study inclusion criteria

We selected articles for the meta-analysis on the basis of the following set of predetermined criteria: (1) Each paper had to report the results of an original research study including an idiom or proverb comprehension or production task, and scores relating to accuracy on the relevant task had to be reported separately from other measures.¹ Studies were excluded if they involved data that were insufficient to calculate effect sizes or if the relevant data could not be obtained from the authors; (2) In each study, participants had to be diagnosed with ASD by experienced clinicians using the standard diagnostic criteria: Diagnostic and Statistical Manual of Mental Disorders (DSM) or International Classification of Diseases (ICD). This included all diagnostic categories of Autism Spectrum Disorder which had previously been considered separate (i.e., Asperger's syndrome and pervasive developmental disorder); (3) The study had to include a TD comparison group, matched to the ASD group on chronological age and IQ.

¹ Two studies in our meta-analysis, reported in Vulchanova et al. (2019), did not meet this strict criterion, as they reported participants' overall scores based on a task consisting of 40 idioms, 16 proverbs and 20 metaphors.

Screening process

The search was conducted by the two authors independently, using three major databases and their search engines. In the first step, we compared the three lists of results (coming from Web of Science, Scopus and ProQuest D&T Global) and removed all duplicates. Then, we read the titles and abstracts of each study. In case the study appeared to meet our eligibility criteria, or if the title and abstract provided insufficient information, the full-text article/dissertation was reviewed. The tripartite inclusion criteria described above were then applied to determine the final selection of papers to be included in the meta-analysis. At each stage, any disagreements between the two raters were resolved by discussion. Details regarding the number of papers considered at each stage of the selection process, as well as the reasons for exclusions, are presented in the PRISMA flow diagram (Moher 2009) in Fig. 1.

Data extraction

For descriptive purposes, we extracted the following study characteristics: title, authors and publication year, language and country, types of tasks involved. The number of ASD and TD participants, and descriptive statistics (means and SDs for performance on the idiom and proverb tasks for both groups) reported in the papers were extracted for the purposes of statistical analyses. In those studies that encompassed neuroimaging or any similar data, we extracted only the behavioural results. When a study included several tasks relating to idiom and proverb processing (or the results were broken down by type of task), the results of these were combined into a single composite measure for the study. This was done because computing effect sizes multiple times based on data coming from the same sample can lead to a misrepresentation of the overall results (see Borenstein et al. 2009). Finally, we extracted the average age and the IQ of the participants in the ASD and TD samples, and recorded the approach used for matching the groups on intellectual ability (i.e., verbal, full-scale or non-verbal IQ).

Meta-analytic procedures and analysis

The statistical analyses were conducted using the Comprehensive Meta-Analysis (CMA) software, version 3 (Biostat). For each effect size a 95% confidence interval was computed, using the original data from the studies. Hedges' g (a variation of Cohen's d correcting for biases caused by small sample sizes; Hedges, 1981; Hedges & Olkin, 1985) was used as our effect size statistic. The Hedges' g value was assigned a positive sign 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 that between-study variations in effect sizes result not only from random error, but also from systematic effects of variables that are likely to vary from study to study (Borenstein et al. 2009).

The possibility that factors beyond an ASD diagnosis had an impact on effect sizes exists because of heterogeneous effect sizes. The heterogeneity of effect sizes was statistically tested using Cochran's Q -statistic. Besides this, we report the I^2 statistic, which expresses the percentage of variation in effect sizes across studies which can be attributed to systematic effects of study variables (rather than chance; Higgins & Thompson, 2002; Higgins et al. 2003). We carried out meta-regression analyses using random-effects models in order to test for the effects of possible moderator variables—the average age and IQ of ASD participants. A funnel plot was created to identify and evaluate the effect of publication bias on the overall effect size estimate. The effect sizes of group differences were plotted on the horizontal axis against the standard error of effect sizes on the vertical axis. The classic fail-Safe N statistic (Rosenthal, 1979) was employed to 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). This statistic 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).

Finally, we have also assessed the quality of the included studies, using a similar procedure to Spain et al. (2018) and Morsanyi et al. (2020b), focusing on the following three criteria: (1) sample characteristics,

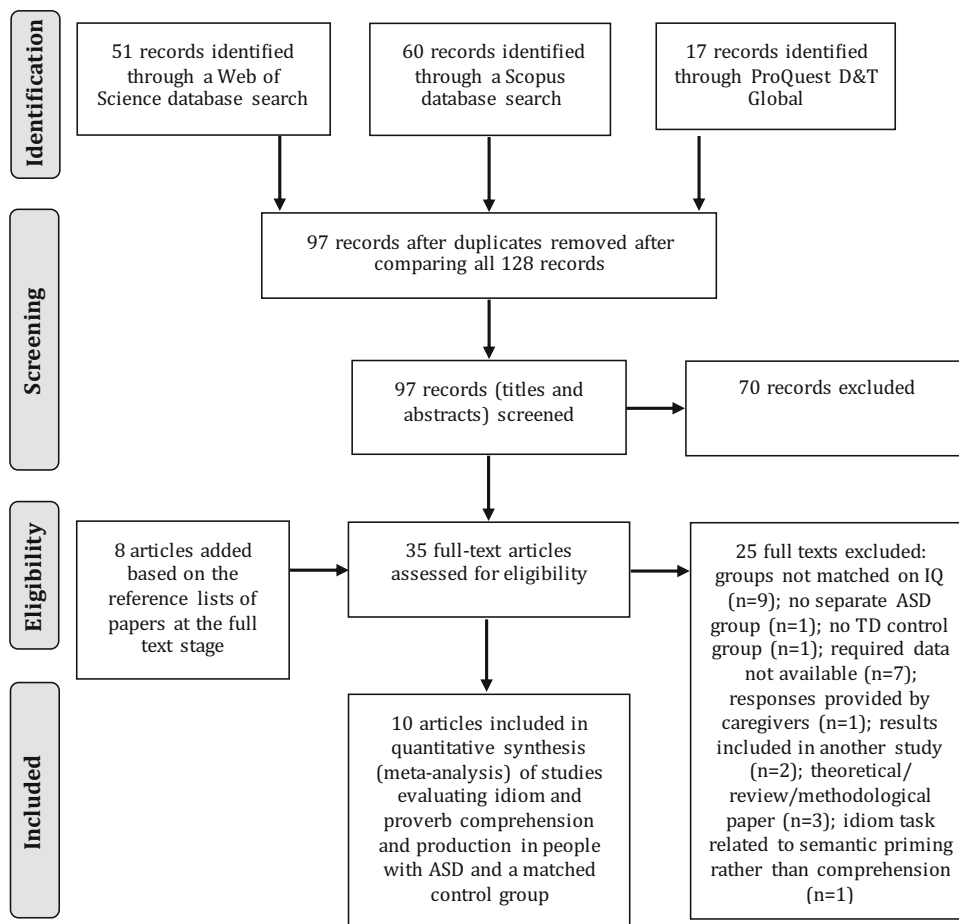


Fig. 1 The PRISMA flow diagram describing the search process, the number of papers included at each step, and the reasons for exclusion at the full-text stage

(2) characteristics of the task, and (3) procedure and materials used for matching the groups on verbal and/or intellectual ability, using the procedures and criteria defined in Morsanyi et al. (2020b).

Results

A total of 11 studies (published in 10 papers), involving 235 ASD and 224 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 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 idiom or proverb task(s), from positive (i.e., the TD group performs better) to negative (i.e., the ASD group performs better). A point estimate of effect sizes is also included in the table. The studies were conducted in five different countries and four different languages (English, Hebrew, Korean, and Spanish), and included participants spanning a reasonably broad age range: from mid-childhood through early adolescence to 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 IQ. The idiom ($n = 6$) and proverb ($n = 4$) tasks used in the studies were diverse in terms of the properties of the idioms

Table 1 Participant characteristics, type of idiom or proverb task, and effect size of group differences in the studies included in the meta-analysis

Publication details, country and language	N ASD (TD)	Mean age ASD (TD)	Mean (V)IQ ASD (TD)	Task type	Point estimate of effect size (Hedges' <i>g</i>)
Tzuriel and Groman (2017) (Israel, Hebrew)	32 (32)	9.33 (9.33)	102 (102)	Proverb explanation (proverb presented in a story context accompanied by a picture); auditory (verbal) presentation	1.26
Saban-Bezael and Mashal (2019) (Israel, Hebrew)	23 (24)	12.7 (12.19)	Only raw scores reported	Idiom comprehension (multiple choice with correct, literal, literal distractor, and unrelated response options); written presentation	1.22
Strandburg et al. (1993) (USA, English)	13 (13)	24.9 (26.2)	95.5 (104.0) ¹	Two-word idiom verification task (including idioms with no literal meaning); written presentation	1.18
Yi et al. (2013) (South Korea, Korean)	12 (12)	8.26 (8.19)	97.17 (100.33)	Conventional and novel proverb comprehension (multiple choice) written presentation	1.12
Mashal and Kasirer (2011) (Israel, Hebrew)	20 (20)	13.02 (not reported)	only raw scores reported	Idiom comprehension (multiple choice with correct, literal, literal distractor, and unrelated response options); written presentation	1.12
Lee et al. (2015) (South Korea, Korean)	16 (10)	9.31 (9.30)	94.75 (not reported)	Idiom comprehension (multiple choice: written idiom presented together with picture representing figurative or literal meaning)	1.10
Whyte et al. (2014) (USA, English)	26 (26)	9.08 (9.03)	120 (127) ²	Idiom explanation (idioms presented in context), auditory (verbal) presentation	0.74
Lee et al. (2016) (South Korea, Korean)	15 (15)	8.34 (8.2)	86.07 (92.27)	Familiar and unfamiliar idiom comprehension (2- or 3-word expressions presented without context; multiple choice with correct, literal and distractor options), written presentation	0.64
Vulchanova et al. (2019) adults (Spain, Spanish)	20 (20)	18.1 (18.3)	122.3 (117.63)	Idiom and proverb comprehension (multiple choice with correct, literal, distractor, and unrelated figurative response options presented as pictures); half of the idioms and proverbs presented auditorily and half presented in a written format	0.33
Vulchanova et al (2019) children (Spain, Spanish)	25 (19)	11.3 (11.9)	118.7 (107.4)	Idiom and proverb verification (multiple choice with correct, literal, distractor, and unrelated figurative response options presented as pictures); half of the idioms and proverbs presented auditorily and half presented in a written format	0.12
McCrimmon et al. (2012) (Canada, English)	33 (33)	18.83 (18.86)	114.1 (109)	Proverb explanation; proverbs presented without context auditory (verbal) presentation	-0.92

The studies are listed in decreasing order of the effect size of group differences

¹Full-scale WAIS IQs for ASD, mean estimated IQ for controls

²The groups were matched on non-verbal IQ

and proverbs, and the modality of presentation, although most studies used a multiple-choice response format.

Figure 2 presents the effect size of the group differences in idiom or proverb comprehension (Hedges' *g* with 95% *CI*s) between individuals with ASD and matched TD controls. Overall, the results

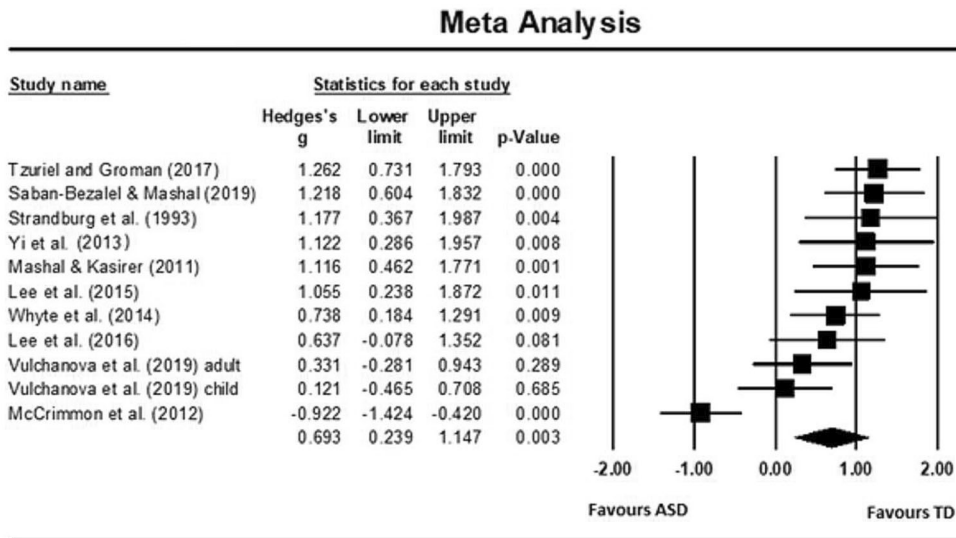


Fig. 2 Hedges' *g* effect sizes with 95% confidence intervals for group differences in idiom and proverb processing accuracy between individuals with ASD and age- and verbal intelligence-

showed a medium sized group difference in idiom or proverb comprehension ($g = 0.69, p < 0.001$; 95% CI 0.24–1.12) in favour of the TD group.² It is notable that there was one study, which showed a large difference in favour of the ASD group (McCrimmon et al. 2012), there were three studies that reported no difference in performance between the groups, and there was also a cluster of 6 studies that reported very similar effect sizes (a large group difference of just above 1 in favour of the TD group). There were two studies (Lee et al. 2015 and Strandburg et al. 1993) that matched the groups on full-scale IQ instead of verbal IQ, and one study (Whyte et al. 2014) that matched the groups on non-verbal IQ. Lee et al. (2015) and Strandburg et al. (1993) were among the cluster of studies that yielded the largest group differences. Leaving out these studies slightly reduced the overall effect size, although it did not have a major effect ($g = 0.60, p < 0.001$; 95% CI 0.004–1.19).

The heterogeneity between studies was significant ($Q(10) = 55.73, p < 0.001, I^2 = 82.05$). This analysis indicated that a considerable proportion of the variance in effect sizes (82%) could be attributed to true

matched controls. The overall mean effect size is presented in the bottom line (and marked by a filled diamond in the figure)

variance rather than to random noise. To investigate the potential causes of variation in effect sizes, we conducted two meta-regression analyses. The first one investigated the relations between participants' age and the magnitude of group differences. This analysis included all studies in our meta-analysis ($n = 11$). The results were non-significant ($p = 0.344$), indicating that participants' age was unrelated to the effect size of group differences. We also performed a meta-regression analysis to investigate the relations between participants' level of (verbal) intelligence and the effect size of group differences. In this analysis, we could only include 9 studies that reported participants' (verbal) intelligence. The model was only marginally significant ($Q(1) = 3.55, p = 0.059$), explaining 36% of the variance in effect sizes between studies. An inspection of the regression plot (Fig. 3) indicated that the mean (verbal) IQ of autistic participants was negatively related to the effect size of group differences (i.e., group differences were smaller in the case of participants with higher verbal ability), although it is worth remembering that this was only a marginally significant trend.

Figure 4 presents a funnel plot which assesses the possible impact of publication bias. This analysis encompassed all studies ($n = 11$) included in our meta-analysis. The funnel plot presents the effect sizes of group difference on the horizontal axis and the

² The interpretation of Hedges' *g* is very similar to Cohen's *d*. The current results suggest that the TD group's performance is .24–1.15 *SD* higher than the ASD group's performance in idiom and proverb processing tasks.

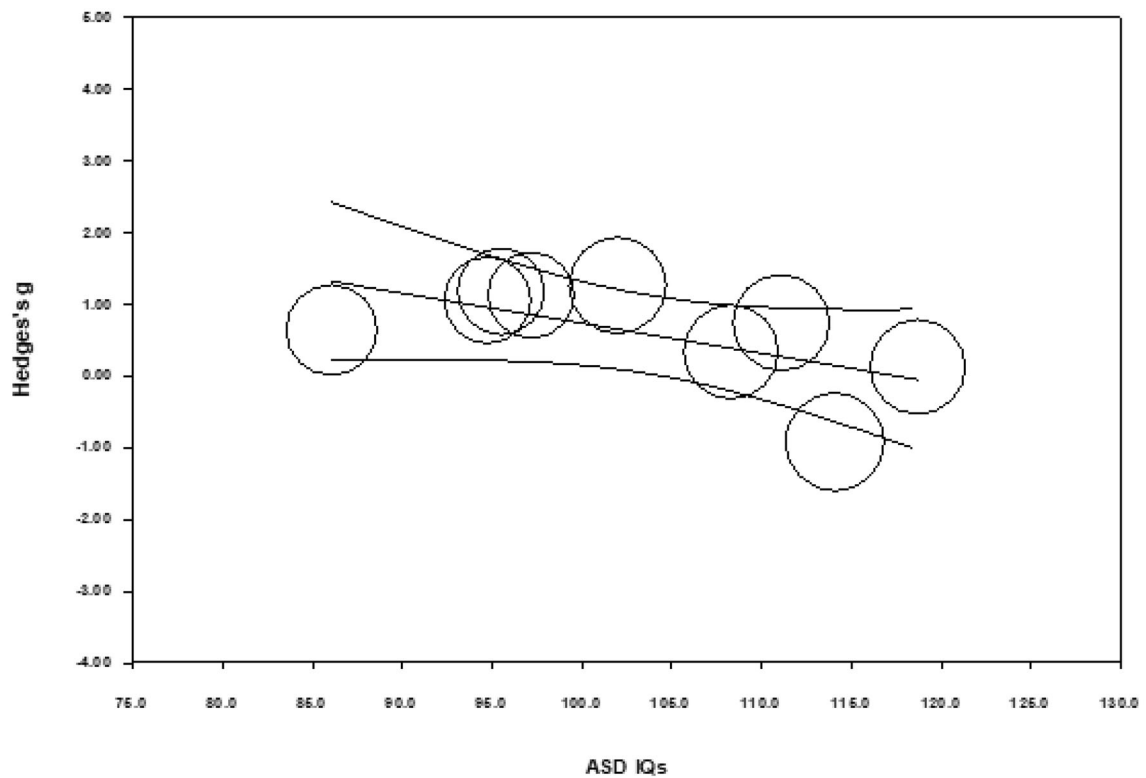


Fig. 3 Plot presenting the regression of Hedges' g effect sizes for group differences in idiom and proverb comprehension scores on the average verbal IQ of the ASD participants in each study

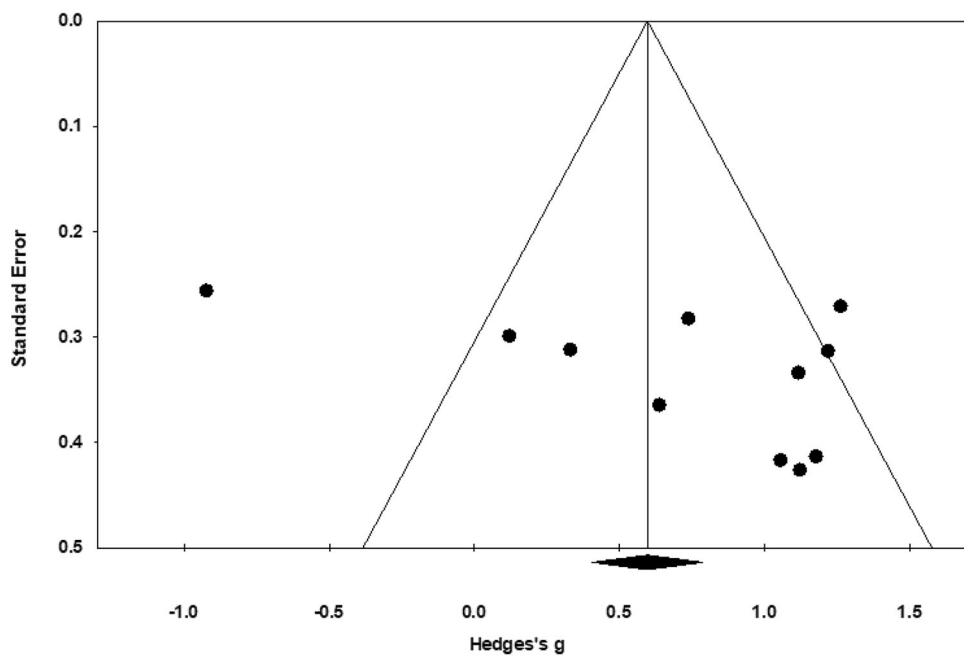


Fig. 4 Funnel plot to assess the potential impact of publication bias

standard error of effect sizes (a sample-size dependent statistic) on the vertical axis. Larger studies (i.e., studies with smaller standard errors) appear toward the top of the graph and, in general, they are expected to cluster around the mean effect size. Smaller studies (i.e., studies with larger standard errors) appear toward the bottom, and are expected to be more dispersed. In the absence of publication bias, studies should be symmetrically distributed on each side of the overall mean effect size (Cooper et al. 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. A visual inspection of the plot showed some asymmetry, with the McCrimmon et al. (2012) study appearing as an outlier. Nevertheless, in the case of a random effects model, it can be difficult to interpret the funnel plot visually (Lau et al. 2006). For this reason, we have also conducted some statistical analyses relating to the potential presence of publication bias. Using the fail-Safe N statistic, our analysis revealed that 117 missing studies with a null finding would be needed to change the overall result of the meta-analysis to not significantly different from zero. 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. In our case, the critical number of missing studies would have to be above 65 for the results to be considered stable. As 117 is well above this critical value, the results can be considered stable (i.e., there is no evidence of a publication bias). This method is, however, less reliable when the number of studies included in the analysis is low (Lau et al. 2006). As an additional check for the possible presence of publication bias, we also computed the correlation between sample sizes and effect sizes. In this analysis we considered McCrimmon et al. (2012) with a positive sign, as our interest was in determining if larger studies tended to yield smaller effect sizes than smaller studies. The correlation was weak, negative and non-significant ($r(8) = -0.07$, $p = 0.839$), indicating that smaller studies were not systematically associated with larger effect sizes.

In terms of the quality of the papers (Table 2), all of them were published in peer-reviewed journals, and we judged that all of them were of at least moderate quality, which suggests that in general the methodology of the papers can be considered robust. The studies

usually matched the groups at least on age and gender (apart from intelligence), and several papers used additional measures as well. About half of the studies gave detailed explanations regarding the inclusion and exclusion criteria that they used. We also listed the specific diagnostic categories of the participants in Table 2 (although in the DSM-5, all of these categories have been merged into a single autism spectrum disorder diagnostic category). Most studies used idiom and proverb tasks that were created by the authors, but these tasks were typically piloted with independent samples, indicating that they were appropriate for the relevant age- and ability-level groups. Although the studies did not generally report the reliability of the idiom and proverb tasks, the number of items tended to be reasonably large. All studies matched the participants on a standardized test of verbal ability or intelligence (usually using a version of the Wechsler intelligence scale), although the details of the instrument used by Lee et al. (2015) were not described. Some studies assessed language skills with multiple tests, although the most common strategy was to match the groups on the Vocabulary subtest of the Wechsler Intelligence Scale.

Discussion

This meta-analysis focused on idioms and proverbs, two special forms of figurative language, which are more opaque and conventionalized than metaphors, and, for this reason, rely more heavily on the processing of contextual cues, and might also be more difficult to process (cf., Vulchanova et al. 2015). Idioms and proverbs are also likely to be more language-related than metaphors, the processing of which, in some contemporary views, go beyond language (see Holyoak, 2019; Holyoak & Stamenković, 2018).

Our analysis revealed a medium size group difference in processing idioms and proverbs in autism, with a point estimate of $g = 0.69$. A detailed evaluation of the methods presented in the papers suggested that the quality of the studies was generally good. Moreover, there was no evidence of publication bias, which suggests that our estimate of the overall effect size may be considered reliable. The overall effect size in our study is very similar to the effect size of group differences in the case of metaphors reported in recent

Table 2 Sample characteristics, design and properties of the idiom and proverb tasks, materials used for matching the groups on (verbal) IQ, and publication status of the studies included in the meta-analysis, with global quality assessment

Publication details	Sample characteristics	Characteristics of the idiom or proverb tasks	Materials used for matching the groups on IQ/verbal ability	Publication status	Global quality rating
Tzuriel and Groman (2017)	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 socio-economic status	Children's Proverbial Understanding Test (Tzuriel and Valdman 2009): Understanding the proverb and its context subtest, 11 items	WISC-VI Vocabulary subtest	Journal paper	Strong
Saban-Bezalel and Mashal (2019)	Participants recruited from schools. ASD group diagnosed using DSM-IV-TR or DSM-5 criteria, included pervasive developmental disorder, autism, high functioning ASD and ASD, verified by Social Communication Questionnaire, matched on age, gender executive functioning and vocabulary abilities. Exclusions or withdrawals not described	20-item idiom comprehension test based on Mashal et al. (2008)	WISC-IV HEB (Wechsler 2003) Vocabulary subtest	Journal paper	Strong
Strandburg et al. (1993)	Adults meeting DSM III criteria for autism and TD matched on age, gender and full-scale IQ; ASD group recruited from community, 2 participants excluded due to comorbid psychiatric conditions	Verification task created by the authors including 40 idiomatic expressions, piloted in an independent sample	Full-scale WAIS for ASD, WAIS Vocabulary and Block Design subtests for TD	Journal paper	Strong
Yi et al. (2013) (South Korea, Korean)	Asperger's syndrome; recruitment process and exclusions or withdrawals not described; groups matched on age, gender, VIQ	Task created by the authors: transparent and opaque proverbs with prior context, piloted in independent samples	Receptive and expressive vocabulary test (Kim et al. 2009), Korean sentence comprehension test (Pae et al. 2004), revised Chinese version of WISC verbal IQ	Journal paper	Moderate
Mashal and Kasirer (2011)	ASD participants diagnosed on the basis of DSM-IV-TR, recruited from special classes for autistic children within mainstream schools; inclusion criteria: speaks fluently, monolingual Hebrew home environment, no neurological impairment, scored within the normal range on an abstract verbal reasoning test;	20-item idiom comprehension test based on Mashal et al. (2008)	Reading accuracy and fluency test created by the authors; synonyms subtest of the MEM (a Hebrew abstract verbal reasoning test; Glantz, 2008)	Journal paper	Moderate to strong

Table 2 continued

Publication details	Sample characteristics	Characteristics of the idiom or proverb tasks	Materials used for matching the groups on IQ/verbal ability	Publication status	Global quality rating
	matched to TD on reading accuracy and fluency				
Lee et al. (2015)	Patients at a child and adolescent psychiatry clinic diagnosed using DSM-IV-TR criteria, matched to TD on age, gender and IQ; exclusions not described	Task created by the authors: 15 idioms presented with picture representing literal meaning (Cronbach's alpha: 0.94) and 15 idioms presented with picture representing figurative meaning (Cronbach's alpha: 0.79)	No test specified	Journal paper	Moderate
Whyte et al. (2014)	ASD group based on parental report of a previous diagnosis of autism spectrum disorder, groups matched on chronological age and nonverbal IQ	Task created by the authors: contextualized idioms ($n = 20$) selected from Titone and Connine (1994)	NVIQ raw score and age-equivalent score (based on the Matrices subtest of the KBIT-2; Kaufman and Kaufman 2004)	Journal paper	Moderate
Lee et al. (2016)	High-functioning ASD group included children diagnosed with ASD or pervasive developmental disorder using DSM-IV criteria, children were only included if their receptive and expressive vocabulary and sentence comprehension scores were in the normal range, and they had no vision or hearing problems. Groups well-matched on age, and sentence comprehension (marginal differences in performance IQ, expressive and receptive vocabulary)	Task created by the authors: 40 idioms low or high in familiarity representing actions or emotions	Receptive and Expressive Vocabulary Test (REVT; Kim et al. 2009), Korean Oral Syntax Expression Comprehension Test (KOSECT; Pae et al. 2004), K-WISC IV (performance IQ)	Journal paper	Moderate
Vulchanova et al. (2019) adults	High-functioning autism (confirmed with ADOS); recruitment details and exclusions or withdrawals not described; groups matched on age, gender, full-scale IQ and verbal comprehension	Task created by the authors: 20 biological idioms, 20 cultural idioms and 16 proverbs (mixed with 20 metaphors); piloted with independent samples	Full-scale IQ and verbal comprehension based on the WAIS-IV	Journal paper	Moderate to strong
Vulchanova et al (2019) children	High-functioning autism (confirmed with ADOS); recruitment details and exclusions or withdrawals not described; groups matched on age, gender, full-scale IQ and verbal comprehension	Task created by the authors: 20 biological idioms, 20 cultural idioms and 16 proverbs (mixed with 20 metaphors); piloted with independent samples	full-scale IQ and verbal comprehension based on the WISC-IV	Journal paper	Moderate to strong

Table 2 continued

Publication details	Sample characteristics	Characteristics of the idiom or proverb tasks	Materials used for matching the groups on IQ/verbal ability	Publication status	Global quality rating
McCrimmon et al. (2012)	Adolescents and young adults diagnosed with Asperger's syndrome, confirmed using DSM-IV-TR criteria. ASD sample recruited through community support agencies, TD sample recruited through the secondary or post-secondary school system. Inclusion criteria for ASD group: intact language development in childhood, IQ of 85 or higher. Exclusions: 6 individuals were excluded, because they had a diagnosis of autistic disorder (not Asperger's), 2 individuals excluded, because they did not meet the IQ criterion. TD participants had to have no history of mental health problems. Groups matched on performance and full-scale IQ, marginal difference in verbal IQ	Proverb subtest of the Delis–Kaplan Executive Function System (DKEFS) (Delis et al. 2001): 8 proverbs (common and uncommon)	Full-scale IQ and performance IQ (WASI, 1999); there was a marginal difference in verbal IQ in favour of the Asperger's group	Journal paper	strong

meta-analyses ($g = 0.72$ in Kalandadze et al. 2018, and $g = 0.76$ in Morsanyi et al. 2020b). This suggests that although the cognitive processes involved in understanding metaphors might be different from how idioms and proverbs are processed, and idioms and proverbs might be more bound to verbal skills, the difference in their levels of difficulty is not very substantial. Indeed, the heterogeneity of group differences reported by the individual studies (ranging from non-significant to large in size, and including differences in favour of both the TD and the ASD groups) suggests that factors other than the type of figurative language (such as the characteristics of the participants involved in the studies, and the properties of the idiom and proverb processing tasks) might be responsible for much of the variance.

In our analyses, we only included studies where the autistic and TD samples were matched on both age and

(preferably verbal) intelligence. Several theorists proposed that impairments in figurative language are related to more general issues with language skills in autism (e.g., Brock et al. 2008; Gernsbacher & Pripas-Kapit, 2012; Geurts et al. 2020; Norbury, 2005), and when the groups are well-matched in terms of their language skills, group differences should disappear. There were two studies (Lee et al. 2015 and Strandburg et al. 1993) that matched the groups on full-scale IQ instead of verbal IQ, and one study (Whyte et al. 2014) that matched the groups on non-verbal intelligence. The effect sizes reported by these studies were among the largest, indicating that matching strategy might impact the size of group differences. Nevertheless, excluding these studies only reduced slightly the overall effect size estimate (from 0.69 to 0.60), which suggests that matching strategy (in terms of whether

verbal, non-verbal or full-scale IQ was used) did not have a very considerable effect on the results.

We also conducted a meta-regression analysis to investigate the potential relation between participants' level of intelligence and the effect size of group differences. This analysis yielded a marginally significant result, indicating a trend for smaller group differences in the case of participants with higher (verbal) ability levels. This suggests that group differences might be reduced or absent in the case of participants with particularly high verbal abilities. Although this result was not statistically significant (most likely due to a lack of statistical power, as we could only include 9 studies in this analysis), there are reasons to believe that this finding might be meaningful. One reason is that the effect size relating to the effect of (verbal) ability on the size of group differences was moderate ($R^2 = 0.36$), and because this result replicated the findings of Morsanyi et al. (2020b) in relation to metaphor processing. These results suggest that closely matching the autistic and TD groups on verbal skills might not be enough to eliminate group differences, unless participants have very high verbal ability. This finding contrasts with claims in the literature that matching ASD and TD participants closely on verbal ability eliminates group differences (e.g., Brock et al. 2008; Gernsbacher & Pripas-Kapit, 2012; Geurts et al. 2020; Norbury, 2005), although we should note that most studies did not use an extensive range of tasks to match the groups on verbal ability.

One study included in our meta-analysis (McCrimmon et al. 2012) reported a large group difference in proverb explanation in favour of the ASD group (comprising of participants diagnosed with Asperger's syndrome, with no history of a language delay). The findings of this study stand in contrast with the other studies in our meta-analysis, which all reported a group difference in favour of the TD participants, although this difference was negligible in some cases. The findings of this study are even more surprising, given that previous reviews suggested that verbally explaining the meaning of figures of speech is particularly challenging for autistic participants (cf., Kalandadze et al. 2019). A potential reason for this seemingly anomalous finding is that, as idioms and proverbs are conventionalized, and their representations are stored in the mental lexicon (Conklin & Schmitt, 2012), explaining their meaning is not too

dissimilar to explaining the meaning of individual words, which tends to be a peak ability in Asperger's syndrome (see Soulières et al. 2011). In contrast to McCrimmon et al. (2012), Tzuriel and Groman (2017) found a large difference in favour of the TD group on a proverb explanation task (see Fig. 2). Although, both in McCrimmon et al. (2012) and Tzuriel and Groman (2017) the materials were presented verbally by the experimenter, it is likely that the latter study was more similar to an everyday conversational situation, as the materials required taking into account a story context. This could have increased the processing demands of the task in the case of the ASD group.

The finding that participants with ASD do not always have problems with figurative language processing, and sometimes even outperform TD participants, fits well with suggestions that pragmatic impairments in autism are neither global nor uniform (Geurts et al. 2020), and also with recent changes in the DSM diagnostic criteria for autism that no longer involve language impairments (although impairments of non-verbal communication still feature prominently). Indeed, overall, our findings suggest that autistic participants with high verbal ability are able to perform idiom and proverb processing tasks with a similar level of accuracy as age- and verbal intelligence-matched TD participants (and sometimes even better).

Relating to the above point, it should be noted that an absence of group differences in accuracy does not necessarily imply an absence of differences in the strategies that participants use when processing idioms and proverbs. For example, Vulchanova et al. (2019) reported no group differences in accuracy on a sentence-picture matching task involving idioms and proverbs either in the case of children or adults. Nevertheless, eye-tracking and mouse-tracking results showed that participants with autism spent more time on considering the literal (as well as the figurative) interpretations of the idioms and proverbs presented to them. This might suggest that they were less able to suppress a literal interpretation of these figures of speech (see also Gold & Faust, 2010; Melogno et al. 2017; Vulchanova et al. 2019), or could simply point at a processing style where participants with autism take longer to make decisions and avoid "jumping to conclusions" (Brosnan et al. 2014).

Saban-Bezael and Mashal (2015) also reported no difference in the performance of autistic and non-

autistic participants on a lexical decision task related to idiom processing. However, they did find evidence for differences in the hemispheric processing of these expressions (with a right hemisphere advantage in the TD group, but bilateral processing in the ASD group). A potential explanation is that the ASD group relied on a compensatory mechanism. If this is the case, this might explain why ASD participants with higher levels of intelligence are more likely to show no impairments in figurative language processing.

In a separate meta-regression analysis, we also investigated the relation between participants' age and the effect size of group differences. In this analysis, we were able to include all studies from our meta-analysis. There are suggestions in the literature that there is a developmental delay in autism in figurative language processing (cf., Saban-Bezeal & Mashal, 2018; Vulchanova et al. 2015). If this was the case, we could expect that the effect size of group differences might decrease with age. Idiom processing shows a protracted development from around the age of five to adolescence (see Hattouti et al. 2016 for a review). The samples included in our meta-regression covered a broad age range from mid-childhood to young adulthood, which made it possible to look at developmental differences at the time period which is critical for the development of idiom and proverb processing. The meta-regression relating to the effect of participants' age yielded a non-significant result, which suggested that the effect size of group differences neither increased nor decreased with age. This finding speaks against a general developmental delay in idiom and proverb processing in autism.

Limitations

The most important limitation of this study is the relatively small number of studies that we were able to include in our analyses. Although our overall estimate for the average effect size of group differences is likely to be reasonably accurate, we experienced a problem with statistical power in the case of our meta-regression analysis relating to the effects of (verbal) ability. Another factor that limits our ability to advance current debates in the field, relating to the role of specific language skills in figurative language in autism, is that most studies in our review used a vocabulary test to match the groups. This subtest is considered to provide a good measure of an

individual's expressive vocabulary, verbal knowledge, and crystallized and general intelligence, and it also draws heavily on memory, learning ability, and concept- and language development (Sattler, 1988). Nevertheless, it does not provide a good measure of syntax, which is considered to play an important role in the processing of idioms, at least in real-life conversational situations (cf., Whyte et al. 2014). Relating to this point, it is notable that most studies in our analysis presented short idioms and proverbs in a written format on a computer screen, without the requirement to process contextual information, which is very different from how idioms and proverbs are typically encountered in everyday situations (where pragmatics play a more important role).

Future directions

Our study suggests that, in general, there is a medium size difference in idiom and proverb processing between autistic and TD participants, in favour of the TD group. Nevertheless, the studies included in our meta-analysis yielded extremely heterogeneous results, which calls for further investigations into the circumstances in which autistic people might perform well or where they might struggle. Given that there was only one study which found an advantage in favour of the ASD group, it would be important to follow up on this finding, and see if it replicates in a more diverse ASD group, also including participants who experienced a delay in their language development.

Another interesting future avenue could be to examine the role of context in understanding idioms and proverbs in autism. So far, no study has manipulated the presence/absence of context. The presence of supporting context facilitates the processing of the figurative meaning of idioms and proverbs in the case of TD individuals (e.g., Vulchanova et al. 2015), helping them to avoid a literal interpretation. Nevertheless, contextual cues might be less helpful for autistic individuals when they have to suspend common interpretation of statements (e.g., Happé, 1997; Joliffe and Baron-Cohen 1999; López & Leekam, 2003; Morsanyi & Handley, 2012—although see e.g., Giora et al. 2012). The investigation of contextual cues could involve the presentation of proverbs and idioms embedded in a text or a story, compared to a decontextualized presentation format. Additionally,

the role of ecologically valid contextual cues (such as prosody and facial expressions) could also be investigated. For example, it has been noted that the auditory presentation format might be difficult for autistic participants (Vulchanova et al. 2019). Nevertheless, within the auditory modality there is likely to be important differences between an experimenter presenting participants with idioms and proverbs, as opposed to an auditory presentation of the same materials by a robotic voice. The inclusion of context would make it possible to evaluate the role of pragmatic abilities in both idiom and proverb comprehension in autism. The results of these investigations may also contribute to the debate related to the competing models of idiom comprehension.

Our decision to consider idioms and proverbs together due to their shared features of conventionality and opacity, and their differences in comparison to metaphor, shall not be interpreted as a claim that the processing of these two forms of figurative language is exactly the same. Nevertheless, we expected that the patterns of impairment in ASD might be similar in the case of idioms and proverbs. Indeed, when we consider the effect sizes of group differences between ASD and TD participants in the case of studies that included idioms vs. proverbs (or both), we can see that the results were overlapping, without an apparent separation between the two types of figurative language. Evaluating differences in idiom and proverb comprehension in both ASD and TD populations might be an interesting future direction in figurative language research.

Another future direction could be to extend the systematic investigation of figurative language processing in autism to other figures of speech. Recent systematic reviews and meta-analyses have investigated figurative language processing in autism in general (Kalandadze et al. 2018), as well as the processing of metaphors (Kalandadze et al. 2019; Morsanyi et al. 2020a, 2020b). An interesting future direction could be to review the studies related to irony—a figure of speech which not only communicates the opposite of what is said, but is also more dependent on pragmatics than metaphor, idioms or proverbs. This is especially true for sarcasm, a subtype of irony which is directed at a person, with the intent to criticize. These phenomena depend on both relevant background knowledge (which allows us to see how an ironic expression contradicts what is usual or

expected) and prosodic features/intonation patterns which make them easier to understand.

Acknowledgements Preparation of this paper was supported by the Ministry of Education, Science and Technological Development of the Republic of Serbia grant to Dušan Stamenković (Contract No. 451-03-9/2021-14/200165).

Availability of data and material This study has been preregistered: https://www.crd.york.ac.uk/PROSPERO/display_record.php?RecordID=235762. Raw data used for the analyses is available from the authors on request. A detailed description of the analysis procedures can be found in the paper.

Declarations

Conflict of interest All authors declares that they have no conflict of interest.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

References

- American Psychiatric Association. (2013). *Diagnostic and statistical manual of mental disorders* (5th ed.). Arlington, VA: Author.
- Bal, V. H., Kim, S.-H., Fok, M., & Lord, C. (2019). Autism spectrum disorder symptoms from ages 2 to 19 years: Implications for diagnosing adolescents and young adults. *Autism Research*, 12, 89–99. <https://doi.org/10.1002/aur.2004>
- Baron-Cohen, S., Tager-Flusberg, H., & Cohen, D. J. (Eds.). (2000). *Understanding other minds: Perspectives from developmental cognitive neuroscience* (2nd ed.). Oxford University Press.
- Bobrow, S. A., & Bell, S. M. (1973). On catching on to idiomatic expressions. *Memory & Cognition*, 1(3), 343–346. <https://doi.org/10.3758/BF03198118>
- Borenstein, M., Hedges, L. V., Higgins, J., & Rothstein, H. R. (2009). *Introduction to meta-analysis*. Wiley. <https://doi.org/10.1002/9780470743386>
- 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>
- Cacciari, C., & Glucksberg, S. (1991). Understanding idiomatic expressions: The contribution of word meanings. In G. B. Simpson (Ed.), *Understanding word and sentence* (pp. 251–273). North-Holland: Elsevier.
- Cacciari, C., & Tabossi, P. (1988). The comprehension of idioms. *Journal of Memory and Language*, 27, 668–683. [https://doi.org/10.1016/0749-596X\(88\)90014-9](https://doi.org/10.1016/0749-596X(88)90014-9)
- 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, e0168571. <https://doi.org/10.1371/journal.pone.0168571>
- Chomsky, N. (1980). *Rules and representations*. New York: Columbia University Press.
- Conklin, K., & Schmitt, N. (2012). The processing of formulaic language. *Annual Review of Applied Linguistics*, 32, 45. <https://doi.org/10.1017/S0267190512000074>
- 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>
- Delis, D., Kaplan, E., & Kramer, J. (2001). *Delis–Kaplan Executive Function System (D-KEFS): Examiner’s manual*. San Antonio, TX: The Psychological Corporation.
- 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>
- Frith, U. (1989). A new look at language and communication in autism. *International Journal of Language & Communication Disorders*, 24, 123–150. <https://doi.org/10.3109/13682828909011952>
- Gentner, D. (2010). Bootstrapping the mind: Analogical processes and symbol systems. *Cognitive Science*, 34, 752–775. <https://doi.org/10.1111/j.1551-6709.2010.01114.x>
- 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.
- Gibbs, R. W., Jr., & Beitel, D. (1995). What proverb understanding reveals about how people think. *Psychological Bulletin*, 118(1), 133–154. <https://doi.org/10.1037/0033-2909.118.1.133>
- Gildea, P., & Glucksberg, S. (1983). On understanding metaphor: The role of context. *Journal of Verbal Learning and Verbal Behavior*, 22, 577–590. [https://doi.org/10.1016/S0022-5371\(83\)90355-9](https://doi.org/10.1016/S0022-5371(83)90355-9)
- Giora, R., Gazal, O., Goldstein, I., Fein, O., & Stringaris, A. (2012). Salience and context: Interpretation of metaphorical and literal language by young adults diagnosed with Asperger’s syndrome. *Metaphor and Symbol*, 27(1), 22–54. <https://doi.org/10.1080/10926488.2012.638823>
- Glantz, I. (2008). *Hemed* (8th ed.). Barak Information Systems.
- 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>
- Hamblin, J. L., & Gibbs, R. W., Jr. (1999). Why you can’t kick the bucket as you slowly die: Verbs in idiom comprehension. *Journal of Psycholinguistic Research*, 28(1), 25–39. <https://doi.org/10.1023/A:1023235403250>
- 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. G. (1995). Understanding minds and metaphors: Insights from the study of figurative language in autism. *Metaphor and Symbol*, 10(4), 275–295.
- Happé, F. G. (1997). Central coherence and theory of mind in autism: Reading homographs in context. *British Journal of Developmental Psychology*, 15(1), 1–12. <https://doi.org/10.1111/j.2044-835X.1997.tb00721.x>
- 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>
- Hattouti, J., Gil, S., & Laval, V. (2016). Le développement de la compréhension des expressions idiomatiques: une revue de littérature. *L’année Psychologique*, 1(1), 105–136. <https://doi.org/10.4074/S0003503316000294>
- 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*. Cambridge: 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>
- 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). Oxford: Oxford University Press.
- Holyoak, K. J. (2019). *The spider’s thread: Metaphor in mind, brain, and poetry*. Cambridge: MIT Press. <https://doi.org/10.7551/mitpress/11119.001.0001>
- Holyoak, K. J., & Lu, H. (2021). Emergence of relational reasoning. *Current Opinion in Behavioral Sciences*, 37, 118–124. <https://doi.org/10.1016/j.cobeha.2020.11.012>

- 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>
- Jarrold, C., & Brock, J. (2004). To match or not to match? Methodological issues in autism-related research. *Journal of Autism and Developmental Disorders*, 34, 81–86. <https://doi.org/10.1023/B:JADD.0000018078.82542.ab>
- Jolliffe, T., & Baron-Cohen, S. (1999). A test of central coherence theory: linguistic processing in high-functioning adults with autism or Asperger syndrome: Is local coherence impaired? *Cognition*, 71(2), 149–185. [https://doi.org/10.1016/S0010-0277\(99\)00022-0](https://doi.org/10.1016/S0010-0277(99)00022-0)
- Kalandadze, T., Norbury, C., Nærland, T., & Næss, K.-A.B. (2018). Figurative language comprehension in individuals with autism spectrum disorder: A meta-analytic review. *Autism*, 22, 99–117. <https://doi.org/10.1177/1362361316668652>
- Kalandadze, T., Bambini, 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>
- Kanner, L. (1943). Autistic disturbances of affective contact. *Nervous Child*, 2, 217–250.
- Kaufman, A. S., & Kaufman, N. L. (2004). *Kaufman Brief Intelligence Test* (2nd ed.). Bloomington, MN: Pearson Inc.
- Kim, Y. T., Hong, G. H., Kim, K. H., Jang, H. S., & Lee, J. Y. (2009). *Receptive & expressive vocabulary test (REVT)*. Seoul Community Rehabilitation Center.
- Kintsch, W. (2000). Metaphor comprehension: A computational theory. *Psychonomic Bulletin & Review*, 7, 257–266. <https://doi.org/10.3758/BF03212981>
- 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>
- Lee, S. B., Song, S. H., Ham, J. H., Song, D. H., & Cheon, K.-A. (2015). Idiom comprehension deficits in high-functioning autism spectrum disorder using a Korean autism social language task. *Yonsei Medical Journal*, 56(6), 1613–1618. <https://doi.org/10.3349/ymj.2015.56.6.1613>
- Lee, S.-I., Bae, H.-S., & Lee, Y. (2016). Comprehension abilities of idioms according to semantic types and familiarity in school-aged children with high-functioning autism spectrum disorder. *Communication Sciences & Disorders*, 21(2), 230–243. <https://doi.org/10.12963/csd.16313>
- Lobban-Shymko, J., Im-Bolter, N., & Freeman, N. (2017). Early social communicative skills as predictors of symptom severity in autism spectrum disorder. *Autism & Developmental Language Impairments*, 2, 1–11. <https://doi.org/10.1177/2396941517743418>
- 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).
- López, B., & Leekam, S. R. (2003). Do children with autism fail to process information in context? *Journal of Child Psychology and Psychiatry*, 44(2), 285–300. <https://doi.org/10.1111/1469-7610.00121>
- Maenner M. J., Shaw K. A., Baio J., et al. (2020). Prevalence of Autism Spectrum Disorder among children aged 8 years – Autism and Developmental Disabilities Monitoring Network, 11 Sites, United States, 2016. *The Morbidity and Mortality Weekly Report (MMWR) Surveillance Summaries*, 69(No. SS-4), 1–12. <https://doi.org/10.15585/mmwr.ss6904a1>
- Mashal, N., Faust, M., Hendler, T., & Jung-Beeman, M. (2008). Hemispheric differences in processing the literal interpretation of idioms: Converging evidence from behavioral and fMRI studies. *Cortex*, 44(7), 848–860.
- Mashal, N., & Kasirer, A. (2011). Thinking maps enhance metaphoric competence in children with autism and learning disabilities. *Research in Developmental Disabilities*, 32, 2045–2054. <https://doi.org/10.1016/j.ridd.2011.08.012>
- Mashal, N., & Kasirer, A. (2012). Principal component analysis study of visual and verbal metaphoric comprehension in children with autism and learning disabilities. *Research in Developmental Disabilities*, 33(1), 274–282.
- McCrimmon, A. W., Schwean, V. L., Saklofske, D. H., Montgomery, J. M., & Brady, D. I. (2012). Executive functions in Asperger’s syndrome: An empirical investigation of verbal and nonverbal skills. *Research in Autism Spectrum Disorders*, 6(1), 224–233. <https://doi.org/10.1016/j.rasd.2011.05.003>
- Melogno, S., Pinto, M. A., & Orsolini, M. (2017). Novel Metaphors comprehension in a child with high-functioning autism spectrum disorder: A study on assessment and treatment. *Frontiers in Psychology*, 7, 2004. <https://doi.org/10.3389/fpsyg.2016.02004>
- Mieder, W. (2004). *Proverbs: A handbook*. Westport: Greenwood Press.
- 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. <https://doi.org/10.1371/journal.pmed1000097>
- Morsanyi, K., & Handley, S. J. (2012). Reasoning on the basis of fantasy content: Two studies with high-functioning autistic adolescents. *Journal of Autism and Developmental Disorders*, 42(11), 2297–2311. <https://doi.org/10.1007/s10803-012-1477-0>
- 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. S. (2010). Decontextualised minds: Adolescents with autism are less susceptible to the conjunction fallacy than typically developing adolescents. *Journal of Autism and Developmental Disorders*, 40(11), 1378–1388. <https://doi.org/10.1007/s10803-010-0993-z>
- Morsanyi, K., Stamenković, D., & Holyoak, K. J. (2020b). Metaphor processing in autism: A systematic review and meta-analysis. *Developmental Review*, 57, 100925. <https://doi.org/10.1016/j.dr.2020.100925>
- Morsanyi, K., Stamenković, D., & Holyoak, K. J. (2020a). 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>
- 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>
- Ortony, A., Schallert, D. L., Reynolds, R. E., & Antos, S. L. (1978). Interpreting metaphors and idioms: Some effects of context on comprehension. *Journal of Verbal Learning and Verbal Behavior*, 17, 465–477. [https://doi.org/10.1016/S0022-5371\(78\)90283-9](https://doi.org/10.1016/S0022-5371(78)90283-9)
- Pae S. Y., Lim S. S., Lee J. H., & Jang H. S. (2004). *Korea Sentence Comprehension Test (KOSECT)*. 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/S0954579406060056>
- Pijnacker, J., Geurts, B., van Lambalgen, M., Buitelaar, J., & Hagoort, P. (2010). Exceptions and anomalies: An ERP study on context sensitivity in autism. *Neuropsychologia*, 48(10), 2940–2951. <https://doi.org/10.1016/j.neuropsychologia.2010.06.003>
- 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>
- Rajendran, G., & Mitchell, P. (2007). Cognitive theories of autism. *Developmental Review*, 27, 224–260. <https://doi.org/10.1016/j.dr.2007.02.001>
- 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>
- Saban-Bezael, R., & Mashal, N. (2015). Hemispheric processing of idioms and irony in adults with and without pervasive developmental disorder. *Journal of Autism and Developmental Disorders*, 45, 3496–3508. <https://doi.org/10.1007/s10803-015-2496-4>
- Saban-Bezael, 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: Cambridge University Press.
- Saban-Bezael, R., & Mashal, N. (2019). different factors predict idiom comprehension in children and adolescents with ASD and typical development. *Journal of Autism and Developmental Disorders*, 49, 4740–4750. <https://doi.org/10.1007/s10803-019-04193-9>
- Saeed, E. (2016). *Semantics* (4th ed.). London: Wiley Blackwell.
- Sattler, J. M. (1988). *Assessment of children* (3rd ed.). Jerome M. Sattler.
- Soulières, I., Dawson, M., Gernsbacher, M. A., & Mottron, L. (2011). The level and nature of autistic intelligence II: what about Asperger syndrome? *PLoS ONE*, 6(9), e25372. <https://doi.org/10.1371/journal.pone.0025372>
- 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>
- Stamenković, D., Ichien, N., & Holyoak, K. J. (2020). Individual differences in comprehension of contextualized metaphors. *Metaphor & Symbol*, 35, 285–301. <https://doi.org/10.1080/10926488.2020.1821203>
- Strandburg, R. J., Marsh, J. T., Brown, W. S., Asarnow, R. S., Guthrie, D., & Higa, J. (1993). Event-related potentials in high-functioning adult autistics: Linguistic and nonlinguistic visual information processing tasks. *Neuropsychologia*, 31(5), 413–434. [https://doi.org/10.1016/0028-3932\(93\)90058-8](https://doi.org/10.1016/0028-3932(93)90058-8)
- Swinney, D. A., & Cutler, A. (1979). The access and processing of idiomatic expressions. *Journal of Verbal Learning and Verbal Behavior*, 18, 523–534. [https://doi.org/10.1016/S0022-5371\(79\)90284-6](https://doi.org/10.1016/S0022-5371(79)90284-6)
- Titone, D. A., & Connine, C. M. (1994). Descriptive norms for 171 idiomatic expressions: Familiarity, compositionality, predictability, and literality. *Metaphor and Symbol*, 9(4), 247–270.
- Titone, D. A., & Connine, C. M. (1999). On the compositional and noncompositional nature of idiomatic expressions. *Journal of Pragmatics*, 31(12), 1655–1674. [https://doi.org/10.1016/S0378-2166\(99\)00008-9](https://doi.org/10.1016/S0378-2166(99)00008-9)
- 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., & Valdman, S. (2009). *The development of proverbial understanding as a function of analogical, metaphoric, and verbal abilities*. Osnabruck, Germany: International Association for Cognitive Education and Psychology.
- 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>
- Vulchanova, M., Vulchanov, V., & Stankova, M. (2011). Idiom comprehension in the first language: A developmental study. *VIAL: Vigo International Journal of Applied Linguistics*, 8, 141–163.
- 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>
- Vulchanova, M., Chahboun, S., Galindo-Prieto, B., & Vulchanov, V. (2019). Gaze and motor traces of language processing: Evidence from autism spectrum disorders in comparison to typical controls. *Cognitive Neuropsychology*, 36(7–8), 383–409. <https://doi.org/10.1080/02643294.2019.1652155>
- Wechsler, D. (2003). *Wechsler intelligence scale for children: WISC-IVHEB* (4th ed.). San Antonio, TX: Psychological Corporation.
- Whyte, E. M., Nelson, K. E., & Scherfa, K. S. (2014). Idiom, syntax, and advanced Theory of Mind abilities in children

- with autism spectrum disorders. *Journal of Speech, Language, and Hearing Research*, 57(1), 120–130. [https://doi.org/10.1044/1092-4388\(2013/12-0308\)](https://doi.org/10.1044/1092-4388(2013/12-0308))
- Wing, J. K., & Wing, L. (1976). Provision of Services. In L. Wing (Ed.), *Early childhood autism: Clinical, educational and social aspects*. Oxford: Pergamon Press.
- Yi, D., Hwang, M., & Lim, J. A. (2013). Proverb comprehension in children with Asperger's disorder: The role of transparency. *Communication Sciences & Disorders*, 18(3), 288–296. <https://doi.org/10.12963/csd.13034>

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.