Applied Psycholinguistics (2019), **40**, 1421–1454
doi:10.1017/S0142716419000328

CAMBRIDGE UNIVERSITY PRESS

ARTICLE

A systematic review and meta-analysis of studies on metaphor comprehension in individuals with autism spectrum disorder: Do task properties matter?

Tamar Kalandadze[1,2,*], Valentina Bambini[3] and Kari-Anne B. Næss[1]

[1]University of Oslo, Norway, [2]Knowledge Centre for Education, University of Stavanger, Norway and [3]University School for Advanced Studies IUSS Pavia, Italy
*Corresponding author. Email: tamar.kalandadze@isp.uio.no/tamar.kalandadze@uis.no

(Received 18 September 2018; revised 27 June 2019; accepted 03 July 2019)

Abstract

Individuals with autism spectrum disorder (ASD) often experience difficulty in comprehending metaphors compared to individuals with typical development (TD). However, there is a large variation in the results across studies, possibly related to the properties of the metaphor tasks. This preregistered systematic review and meta-analysis (a) explored the properties of the metaphor tasks used in ASD research, and (b) investigated the group difference between individuals with ASD and TD on metaphor comprehension, as well as the relationship between the task properties and any between-study variation. A systematic search was undertaken in seven relevant databases. Fourteen studies fulfilled our predetermined inclusion criteria. Across tasks, we detected four types of response format and a great variety of metaphors in terms of familiarity, syntactic structure, and linguistic context. Individuals with TD outperformed individuals with ASD on metaphor comprehension (Hedges' $g = -0.63$). Verbal explanation response format was utilized in the study showing the largest effect size in the group comparison. However, due to the sparse experimental manipulations, the role of task properties could not be established. Future studies should consider and report task properties to determine their role in metaphor comprehension, and to inform experimental paradigms as well as educational assessment.

Keywords: autism spectrum disorder; experimental pragmatics; figurative language; response format

A metaphor is a paradigmatic type of figurative language involving discrepancy between the encoded, "literal" meaning of words, and their occasion-specific use (Camp, 2009; Carston, 2010). Metaphors can appear in many forms, such as "Sally is a chameleon" or "Your theory is falling apart." Accordingly, different accounts of metaphor comprehension have been proposed (see Bowdle & Gentner, 2005; Gibbs, 2011; Gibbs & Tendahl, 2006; Gluksberg, 2001; Wilson, 2011). Among them, pragmatic accounts (e.g., relevance theory) focus on metaphor

in communication, highlighting the inferential mechanisms that lead to adjusting the linguistically encoded concepts to arrive at the speaker's intended meaning (Sperber & Wilson, 2012). For instance, in "Sally is a chameleon," the adjustment results in the broadening of the concept CHAMELEON to include not only a species of lizard but also individuals with certain psychological features (Carston, 2012). In contrast, cognitive linguistics accounts (e.g., conceptual metaphor theory) emphasize the role of metaphor in thought, seeing it in terms of conceptual mappings across cognitive domains (Gibbs, 2011; Lakoff & Johnson, 1980). The conceptual mappings emerge in our metaphorical use of language, as in "Your theory is falling apart," for the mapping of theories onto physical constructs such as buildings (THEORIES ARE BULDINGS).

Regardless of the theoretical approach, there is an agreement that metaphors are a ubiquitous part of language and appear frequently in everyday communication, school-books, academic texts, literature, and media communications (Golden, 2010; Steen, Dorst, & Hermann, 2010). Hence, difficulty in understanding metaphors may impede social communication, the ability to obtain information, as well as academic attainment.

In individuals with typical development (TD) metaphor comprehension skills mature throughout childhood until adolescence, and it is commonly assumed that the age of 10 represent a crucial moment (Lecce, Ronchi, Del Sette, Bischetti, & Bambini, 2019; Winner, Rosenstiel, & Gardner, 1976). There is, however, also awareness that metaphorical competence is evident earlier, if assessed with age-appropriate tasks (Pouscoulous, 2011, 2014; Vosniadou, Ortony, Reynolds, & Wilson, 1984). In contrast, profound and lasting difficulty in metaphor comprehension has traditionally been considered characteristic for individuals with autism spectrum disorder (ASD; Adachi et al., 2004; Happé, 1993; Rundblad & Annaz, 2010a), a neurodevelopmental condition characterized by impairments in social communication and interaction, as well as restricted and stereotyped behaviors (American Psychiatric Association, 2013). In particular, individuals with ASD have been reported to interpret metaphors literally (Happé, 1993), a phenomenon referred to as the "literality bias" or concretism (see Rossetti, Brambilla, & Papagno, 2018, for explanation of these terms).

However, there is a discrepancy in study findings. For example, several studies show no statistically significant difference between ASD and TD groups in figurative language comprehension, including metaphors (Hermann et al., 2013; Kasirer & Mashal, 2014; Mashal & Kasirer, 2011; Norbury, 2005). These findings indicate that variables other than characteristics intrinsic to ASD may explain the variation in results across studies. Group matching strategy and general language ability have previously been found to explain some of the between-study variance in figurative language comprehension (see Kalandadze, Norbury, Nærland, & Næss, 2018, for a review). Yet, the remaining unexplained variance requires an investigation of additional relevant variables.

In the behavioral and neurological literature in TD and clinical populations, there is an agreement that the ability to understand metaphors hinges on the task properties such as response format (i.e., multiple-choice vs. verbal explanation task), or absence of linguistic context (see Pouscoulous, 2011, 2014, for discussion of studies with TD participants, and Rossetti et al., 2018, for discussion of literature on

https://doi.org/10.1017/S0142716419000328 Published online by Cambridge University Press

schizophrenia). For instance, children with TD show earlier competence in metaphor comprehension when tested with an act-out rather than a verbal explanation task, perhaps due to the differences in linguistic and cognitive demands that verbal and other types of tasks pose (Pouscoulous, 2011). Similarly, response format could explain how individuals with ASD perform on metaphor tasks. For example, individuals with ASD might understand metaphors comparably to individuals with TD but have more difficulties in explaining the meaning verbally due to difficulties with expressive language (Kwok, Brown, Smyth, & Cardy, 2015). The same might be true for other properties of the metaphors (e.g., the amount and type of context available to interpret the expression, or the familiarity of the expression; Pouscoulous, 2011, 2014).

Despite this knowledge, the properties of metaphor comprehension assessment tasks in studies that compare individuals with ASD to individuals with TD have yet to be comprehensively and systematically explored. In addition, the potential interrelationships between the task properties and any between-study variation have not been systematically investigated. Reviews that have been conducted focused on ASD and figurative language in general, rather than on metaphor specifically (Gernsbacher & Pripas-Kapit, 2012; Kalandadze et al., 2018; Melogno, Pinto, & Levi, 2012; Vulchanova, Saldaña, Chahboun, & Vulchanov, 2015). However, the comprehension of metaphor might differ from the comprehension of other figurative language types in several respects (Vulchanova, Milburn, Vulchanov, & Baggio, 2019). For example, the comprehension of irony seems to depend on Theory of Mind (i.e., the ability to attribute one's own mental states and those of others) more than comprehension of a metaphor (Happé, 1993). In addition, metonymy is processed faster than metaphor, probably due to the routinization of metonymic shifts (Bambini, Ghio, Moro, & Schumacher, 2013). Moreover, the majority of the existing reviews utilized a narrative approach (Gernsbacher & Pripas-Kapit, 2012; Melogno, Pinto, et al., 2012; Vulchanova et al., 2015), which differs from our systematic approach in fundamental ways, especially regarding transparency and systematicity of methods used (Borenstein, Hedges, Higgins, & Rothstein, 2009).

Here, we provide a novel and thorough systematic review and meta-analysis of the properties of the metaphor tasks used in ASD research. We quantitatively compared performance on metaphor comprehension tasks between groups of individuals with ASD and TD and investigated the potential role of the task properties in between-study variation.

By systematically summarizing and synthesizing the available research in the field fulfilling certain inclusion criteria, our study provides robust results that will ultimately have implications when designing future figurative language/metaphor comprehension research, for advancing assessment practices as well as for guiding the research-based intervention paradigms for individuals with ASD.

The following sections provide an overview of metaphor task properties that have been identified as critical for metaphor comprehension in TD and clinical populations (e.g., Pouscoulous, 2011; Rossetti et al., 2018). These are (a) response format (e.g., multiple-choice, meaningfulness decision, etc.), and (b) linguistic characteristics (metaphor familiarity, syntactic structure of the metaphor, linguistic context, and stimulus modality).

https://doi.org/10.1017/S0142716419000328 Published online by Cambridge University Press

Response format

Evidently, the different ways of eliciting the responses when measuring metaphor comprehension pose diverse cognitive and linguistic demands. For example, earlier studies that tested metaphor comprehension of young children by asking them to explain or paraphrase a metaphor concluded that metaphor comprehension was not fully acquired until later in development (e.g., Winner et al., 1976; see Winner, 1988, for an overview; see Pouscoulous, 2011, 2014, for discussion). Alternatively, these findings may be explained by other variables such as response format demands (Pouscoulous, 2014). For example, metaphor explanation or justification tasks require a participant to articulate associations between metaphor topic and vehicle (e.g., "sister" and "butterfly" in "My sister is a butterfly"). Therefore, performance also depends on metalinguistic judgment as well as expressive language and executive control skills. In addition, verbal explanation tasks require participants to explain the meaning of a metaphor to another person, and are therefore more socially demanding than written or computer-based tasks. Explanation tasks might also trigger the processing of the other person's reactions indicating whether the message was understood or not, thus engaging social-communication skills. By contrast, multiple-choice tasks do not rely on expressive language or meta-linguistic skills and require minimal social interaction with the examiner. However, multiple-choice tasks might be more costly in terms of need for inhibiting the false alternative(s) and selecting the correct one, as suggested by evidence from patients with brain lesions (Rapp, Felsenheimer, Langohr, & Klupp, 2018). The important role of the response format in metaphor comprehension is also supported by studies explicitly comparing different tasks. For instance, a study by Perlini et al. (2018) showed that only results from verbal explanation (but not multiple-choice) tasks yielded statistically significant difference between patients in the early phases of psychosis and controls. In addition, Arcara et al. (2019) showed that individuals with traumatic brain injury have more difficulties in performing verbal explanation tasks on figurative language (especially proverbs) compared with multiple-choice tasks.

Linguistic characteristics

Here, we present available evidence regarding the role played by different linguistic characteristics of the metaphor: metaphor familiarity, syntactic structure of the metaphor, linguistic context, and stimulus modality.

Metaphor familiarity

Metaphors are often differentiated according to whether they are conventional (i.e., well established and often encountered in a language), or novel (i.e., not familiar, based on creative invention; Bowdle & Gentner, 2005; Rossetti et al., 2018; Varga et al., 2014). For instance, a metaphor like "The sky's scarf is colored" (Melogno, D'Ardia, Pinto, & Levi, 2012) is considered novel, while "There is a flood outside the museum" (Rundblad & Annaz, 2010a), where flood refers to "lots of people," is considered a lexicalized/conventional metaphor. Both behavioral and neuroimaging evidence from different populations suggests different processing

https://doi.org/10.1017/S0142716419000328 Published online by Cambridge University Press

patterns for metaphor familiarity modulation, and, in particular, a facilitation for conventional compared to novel metaphors (Bambini, Gentili, Ricciardi, Bertinetto, & Pietrini, 2011; Blasko & Connine, 1993; Gluksberg, Gildea, & Bookin, 1982; Lee & Dapretto, 2006; Mashal, Faust, Hendler, & Jung-Beeman, 2009; Rapp et al., 2018; Rossetti et al., 2018; Varga et al., 2014). This might be because at least highly conventional metaphors are to be retrieved from the long-term memory where they are stored as learned lexical units, whereas novel metaphors might to a greater degree depend on the pragmatic ability to make context-relevant inferences (see Pouscoulous, 2011, 2014; Wilson & Carston, 2006, for discussions). Conventional metaphors may, therefore, be understood more quickly and with less cognitive effort, whereas the online processing required for novel metaphors could result in longer processing time involving pragmatic ability to a greater extent. Nevertheless, the exact nature of the difference in comprehension of conventional versus novel metaphors is still debated (Cardillo, Watson, Schmidt, Kranjec, & Chatterjee, 2012).

Syntactic structure of the metaphors

Metaphors in the literature and discourse appear in various syntactic structures. For example, nominal metaphors express the metaphoric meaning using a noun (e.g., "Caroline is a princess"; Wilson & Carston, 2006), predicate metaphors use a verb (e.g., "The rumor flew through the office"; Utsumi & Sakamoto, 2011), and adjective metaphors use an adjective (e.g., "sharp tongue"; Kasirer & Mashal, 2014).

The cognitive effort required for the comprehension of the metaphors of different syntactic structure is likely to diverge (Cardillo et al., 2012; Chen, Widick, & Chatterjee, 2008). For instance, understanding nominal metaphors is suggested to entail either comparison (the assumption that metaphors convey similarities between semantically distinct concepts; Gentner, Bowdle, Wolff, & Boronat, 2001), categorization, (the establishment of taxonomic relations between semantically distinct concepts; Gluksberg, 2003), or both comparison and categorization (Bowdle & Gentner, 2005). On the contrary, predicate metaphors may be understood through a process of highlighting core abstract conceptual features of a verb (Chen et al., 2008). Adjective metaphors are also said to be comprehended through categorization (Gluksberg, 2001; Glucksberg & Keysar, 1990) or by a two-stage categorization process (Utsumi & Sakamoto, 2007). This variation resulting from the different syntactic structures of metaphors may impact study outcomes.

Linguistic context

Metaphors in real life are usually encountered in sentences and/or discourse. Therefore, presenting metaphors with little or no context creates an artificial situation and may obscure the individual's ability to interpret a metaphorical expression. A number of studies on figurative language in individuals with TD as well as clinical populations (i.e., schizophrenia) suggest that the presence of a supportive context can significantly facilitate access to nonliteral meaning (Chakrabarty et al., 2014; Pouscoulous, 2011, 2014). In line with this, event-related brain potential studies have shown that, in the earlier phases of processing, higher

https://doi.org/10.1017/S0142716419000328 Published online by Cambridge University Press

integration efforts are required for metaphoric expressions presented in minimal context compared to supportive context (Bambini, Bertini, Schaeken, Stella, & Di Russo, 2016).

Stimulus modality

The mode of the metaphor stimuli (i.e., auditory vs. written/visual) may also impact performance. For example, young children are usually tested with auditory tasks where they listen to the verbal metaphors and instructions because of their not-yet-adequate reading ability to complete written tasks or read instructions. However, it is not entirely clear whether and how the stimulus modality impacts metaphor comprehension in older children. In addition, metaphor tasks often incorporate a picture/image component to facilitate comprehension of verbal metaphor (e.g., in Rundblad & Annaz, 2010b). Evidence from brain damaged patients suggests that right-hemisphere damaged patients performed better on a verbal than on a visuoverbal test relative to the control group of participants without brain damage (Rinaldi, Marangolo, & Baldassarri, 2004). This might be explained by a disadvantage in processing visual information or by the challenges associated with cross-modal processing.

In sum, evidence suggests that task properties are essential to performance on metaphor comprehension tasks. This may give rise to different processing strategies in individuals with TD and ASD and affect statistical differences between clinical and control groups. As the task properties are often associated with changes in behavioral and neural response in processing metaphors, psycho- and neurolinguistic studies are increasingly based on extensive ratings of metaphor materials. To this end, norms have been established offering metaphorical expression characterizations along several linguistic dimensions, such as familiarity, interpretability, naturalness, and imageability (e.g., Bambini, Resta, & Grimaldi, 2014; Cardillo, Schmidt, Kranjec, & Chatterjee, 2010; Cardillo, Watson, & Chatterjee, 2017; Jacobs & Kinder, 2017). These linguistic dimensions, however, are much less established in the literature on metaphor comprehension in ASD.

Metaphor comprehension task properties in studies with participants with ASD

Studies that compare individuals with ASD to individuals with TD on metaphor comprehension have employed a variety of tasks with different properties. For example, both Happé (1993) and Norbury (2005) employed a sentence completion task where the participants were asked to finish each sentence with a word they could choose from a list. Another type of multiple-choice format was used in the study conducted by Adachi et al. (2004). They tested metaphor comprehension by metaphoric scenarios where the children were asked to read the questions silently and choose from the four response options (one correct and three incorrect). In their study, Rundblad and Annaz (2010a) employed a different format, whereby open verbal responses were given in response to short stories that were accompanied by images/pictures to aid comprehension.

https://doi.org/10.1017/S0142716419000328 Published online by Cambridge University Press

These studies yielded distinct results regarding the magnitude of group-level differences in metaphor comprehension between individuals with ASD and controls with TD. In particular, Adachi et al. (2004), Happé (1993), and Rundblad and Annaz (2010a) found significantly lower ability to understand metaphors, whereas Norbury (2005) found no statistically significant difference between language-ability matched groups. Presumably, the open verbal response format used in Rundblad & Annaz's (2010a) study could be more challenging for at least some individuals with ASD with impaired metalinguistic, expressive language or executive function-related skills (Bishop & Norbury, 2005; Kwok et al., 2015; Lewis, Murdoch, & Woodyatt, 2007; Melogno, Pinto, & Levi, 2015).

Furthermore, including pictures in a metaphor task may also influence performance of individuals with ASD. In individuals with TD, including pictures in a metaphor task can be an advantage because visualization can aid comprehension of verbal metaphors. Using visual support properly, for example pictures accompanying verbal instruction to aid comprehension, is generally also encouraged in work with individuals with ASD (e.g., Dettmer, Simpson, Smith Myles, & Granz, 2000; Nelson, McDonnell, Johnston, Crompton, & Nelson, 2007; Rao & Gagie, 2006). There is evidence from a priming study of probable benefit of using pictures over words to access meaning in ASD (Kamio & Toichi, 2000). Nevertheless, it should be noted that a task presented in two modalities may be more challenging for individuals with ASD as they may struggle to switch between visual and auditory information. This can be hypothesized on the basis of studies such as Reed and McCarthy (2011), where individuals with ASD, compared with participants with TD, showed greater difficulty when different modalities were employed than when only one modality was required. However, the individual needs vary (Rao & Gagie, 2006), resulting in some individuals with ASD benefiting most from picture support, while others from the written support.

Certain task properties might be more suitable than others for individuals across the spectrum, given the cognitive and linguistic strengths (i.e., unimpaired rote memory or interest in details) and differences or challenges (i.e., executive functions) often observed in this population. For example, with regard to the metaphor familiarity, individuals with ASD might have more difficulties than individuals with TD in understanding novel metaphors because comprehension of novel metaphors involves pragmatic operations to a greater degree than conventional ones (Pouscoulous, 2011). In particular, by being innovative and occasion-specific, novel metaphors rely on pragmatic inference involving context-specific meaning adjustments (Recanati, 2004; Sperber & Wilson, 2012; Wilson & Carston, 2006), while conventional metaphors should depend less on inferencing and more on lexical knowledge (Pouscoulous, 2014). Nevertheless, because they are likely to be stored in the lexicon and thus linked to vocabulary knowledge, conventional metaphors might also pose problems for individuals on the spectrum (Pouscoulous, 2011). Individuals with ASD have often been shown to have compromised or biased vocabulary (Tager-Flusberg, 1992; Tager-Flusberg et al., 1990). As vocabulary knowledge is closely related to metaphor comprehension in individuals with TD (Nippold, 2016), compromised vocabulary knowledge might be linked to difficulties in metaphor comprehension in individuals with ASD with poorer vocabulary.

https://doi.org/10.1017/S0142716419000328 Published online by Cambridge University Press

Table 1. Examples of metaphor task properties taken from the included studies

	Task property	Options	Example
Linguistic characteristics of the stimuli	Familiarity	Conventional	'There is a *flood* outside the museum' (Rundblad & Annaz, 2010a).
		Novel	'The house has a *hat*' (Melogno, D'Ardia, et al., 2012).
	Linguistic structure	Nominal	'The moon is a *light bulb*' (Melogno, D'Ardia, et al., 2012)'.
		Predicate	'When Taro plays soccer, no one at his school *comes close to him*' (Adachi et al., 2004).
		Adjective–noun pair	'*Sharp* tongue' (Kasirer & Mashal, 2014).
	Linguistic context	Minimal	'The sky's *scarf* is colored' (Melogno, D'Ardia, et al., 2012).
		Discourse	'Stuart works at a museum. The museum is in the middle of town near a big river. It is a small museum and not so many people come to the museum. Stuart's boss wants more people to come to the museum. So Stuart prepares a very special exhibition. Stuart's boss tells lots and lots of people about Stuart's exhibition. It is Monday morning and Stuart is at home. Suddenly, the phone rings; it is Stuart's boss. Stuart's boss says: "You did it Stuart! There is a *flood* outside the museum." Stuart runs to the museum to look. What does Stuart see?' (Rundblad & Annaz, 2010a).
	Modality	Auditory	'Some roads are *ribbons*' (Chouinard & Cummine, 2016). Participants heard a sentence
		Written	'When Taro plays soccer, no one at his school *comes close to him*' (Adachi et al., 2004). Children were asked to read each question silently
		Multimodal	A story for '*KNOWING IS SEEING*' conceptual metaphor: 'Kristin is trying to make cookies. She doesn't know how to make them. Her cookie dough looks wrong. Her mom teachers her how to make the dough. Kristin says, "Now I *view* it!"'* (Olofson et al., 2014). After hearing (*auditory*) the metaphoric utterance participants were asked which picture (*pictorial*) displayed the meaning of the target utterance. At the same time, a question mark was displayed on the screen.

https://doi.org/10.1017/S0142716419000328 Published online by Cambridge University Press

Table 1. (*Continued*)

Task property	Options	Example
Response format	Multiple choice	'When Taro plays soccer, no one at his school *comes close to him*' (Adachi et al., 2004). Choices: 'Taro ... Is the best soccer player in the school Is the worst soccer player in the school Is sitting to the right of all the students Taro thought he was going to play soccer Don't know' Children were asked to read each question silently
	Verbal explanation	Instructions to the participants: "When I say, 'Is the sun a ball'? you will tell me what you think that means. Or when I say, 'Does the sun have arms?' then you will say what this means to you. Right, now it's your turn." 'The sky's *scarf* is colored' (Melogno, D'Ardia, et al., 2012).
	Multiple-choice combined with verbal justification	Ten metaphors are incorporated into short sentences. The sentences are printed on a card in the Written Metaphor test. Each sentence is followed by three possible responses: the correct metaphorical (the target), the literal (a concrete type of incorrect choice), and the inappropriate meaning (another type of incorrect choice). The participant is asked to listen to the metaphorical sentence and then to point to the one that explains it. After making the choice, each participant is requested to give his/her own interpretation of the metaphor (Gunter et al., 2002).
	Meaningful-ness decision	Participants were instructed to read the words, and indicate as rapidly and accurately as possible whether the target expression was meaningful by lifting and moving the right index finger from the middle position to the right or left key (Gold & Faust, 2010).

Note: Original items extracted from the studies are enclosed within single quotation marks with metaphoric vehicles italicized. The instruction directly cited is enclosed within double quotation marks.
*The expression "Now I view it" is used as a novel version of the conceptual mapping KNOWING IS SEEING, as opposed to the lexicalized version "Now I see it."

Some examples of the different task properties employed in the ASD literature on metaphor comprehension are provided in Table 1. The substantial variability in the assessment tasks employed may account for differences in the results of the studies, making it critical to inspect the properties of these tasks. This issue has been highlighted in a few narrative reviews. For instance, Melogno, D'ardia, et al. (2012) stressed the heterogeneity of the tasks requiring diverse comprehension skills

https://doi.org/10.1017/S0142716419000328 Published online by Cambridge University Press

as the main difficulty in assessing the contribution of different tasks/variables, and they emphasized the urgent need of a careful review of the literature. Likewise, a more recent review by Siqueira, Marques, and Gibbs (2016) claimed that contrasting findings across studies of figurative language (including metaphors) in different clinical populations (including ASD) may be related more to issues related to data collection than to a specific difficulty one population may have in understanding a certain type of figurative language.

The current study: objectives and research questions

The overarching aim of this study was to advance the knowledge and awareness of the impact of task properties on metaphor comprehension performance in individuals with ASD compared to individuals with TD. We aimed to accumulate the existing knowledge by synthesizing the earlier research using the methods of systematic review and a meta-analysis.

The present study (a) explored the properties of the metaphor tasks used in ASD research; (b) investigated the group difference between individuals with ASD and TD on metaphor comprehension, as well as the relationship between the task properties and any between-study variation. We anticipated larger between-study differences in studies employing verbal explanation formats than studies using alternative response formats (e.g., multiple-choice response format).

Method

This study was preregistered in the International Register of Systematic Reviews, PROSPERO, with the registration number CRD42017057231 (available from http://www.crd.york.ac.uk/PROSPERO/display_record.php?ID=CRD42017057231). A dual approach was utilized: a systematic review and a meta-analysis. Preferred Reporting Items for Systematic Reviews and Meta-Analysis (PRISMA) statement (http://www.prisma-statement.org/) was consulted to ensure methodological rigor.

We systematically reviewed the included studies in terms of metaphor task properties (response format and linguistic characteristics). We then undertook a meta-analysis to compare individuals with ASD to individuals with TD on metaphor comprehension, as well as to examine the relationship between response format and any between-study variation.

Data collection, study inclusion, and coding

A systematic literature search was initially conducted on April 14, 2016, and was updated on April 4, 2017. The words for the literature search and the search strategies were selected after discussions in the authors' team and in close collaboration with two librarians at the University of Oslo library with expertise in literature searching. The librarians' responsibility was to ensure that the right search strategies were used and adapted correctly to the different databases. The following electronic databases were searched: Psychinfo, Linguistics and Language Behavior Abstracts (LLBA), Eric, Embase, Norart, Medline, Web of science. The following terms were used as keywords: *ASD OR asperger* OR autis* OR "pervasive developmental*

https://doi.org/10.1017/S0142716419000328 Published online by Cambridge University Press

disorder" combined with *allegor* OR analogy OR analogies OR "figure* of speech" OR "figurative language" OR imagery OR imageries OR metaphor* OR simile**. No restrictions in terms of the publication year were applied.

In addition to the searches in the databases, the key terms (*ASD and metaphor comprehension*; *Asperger and metaphor comprehension*) were applied to Google scholar to identify any gray literature (literature that are not published in scientific journals, e.g., working papers, conference proceedings) to minimize potential publication bias in the meta-analysis. This step is important because studies with significant results and large effect sizes are more easily published than studies that report nonsignificant findings or small effect sizes (Borenstein, et al., 2009). In addition, we manually searched the tables of contents of the following key journals: *Journal of Autism and Developmental Disorders* and *Autism*. Finally, we went through the reference lists of the included articles and book chapters.

To be included in both the systematic review and the meta-analysis, articles were required to meet the following predetermined criteria: (a) the studies had to report on metaphor comprehension separately (when results on metaphor comprehension were part of the results on one global figurative language variable the study was excluded); (b) only participants with ASD were included. Of note, although we consistently use the term "ASD" according to the DSM-5 (American Psychiatric Association, 2013), we expected that diagnoses in the included studies would be based on the *Diagnostic and Statistical Manual of Mental Disorders* (DSM-IV or DSM-IV-TR; American Psychiatric Association, 1994, 2000), or *International Classification of Diseases* (ICD-10; World Health Organization, 1992) criteria, which prevailed at the time the studies was conducted. Thus, participants might have been diagnosed with autistic disorder, Asperger's syndrome/disorder, or pervasive developmental disorder—not otherwise specified; (c) only the studies involving participants with primary diagnosis of ASD (without any co-occurring conditions) were included to avoid the influence of other conditions on the outcome; (d) study design had to compare individuals with ASD to individuals with TD (the groups could either be equated for chronological age [CA], CA and other variables including verbal abilities, or verbal abilities only). No CA restrictions were applied because metaphor comprehension difficulties in ASD are also found in adults with ASD (Happé, 1993); (e) studies had to report data necessary to calculate effect sizes such as mean and standard deviation or p values as well as information and/or examples about the metaphor stimuli that were used; (f) studies could be reported in English, Norwegian, Italian, Russian, Swedish, or Danish because at least one of the authors is competent in each of these languages. By including several languages, we aimed to avoid the language bias often observed in systematic reviews, which is characterized by overrepresentation of English studies (Borenstein et al., 2009). Titles and abstracts obtained from the search were screened for relevance based on the predetermined inclusion criteria by the first author. In case of insufficient information to decide the relevance on the study in the title and abstract, the full-text was reviewed. Finally, 14 studies met the inclusion criteria. For further information on the screening process and a summary of the reasons that studies were excluded see Figure 1.

We coded the following study characteristics: author(s), publication year, diagnostic status, comparison group, CA of the participants (mean and standard

https://doi.org/10.1017/S0142716419000328 Published online by Cambridge University Press

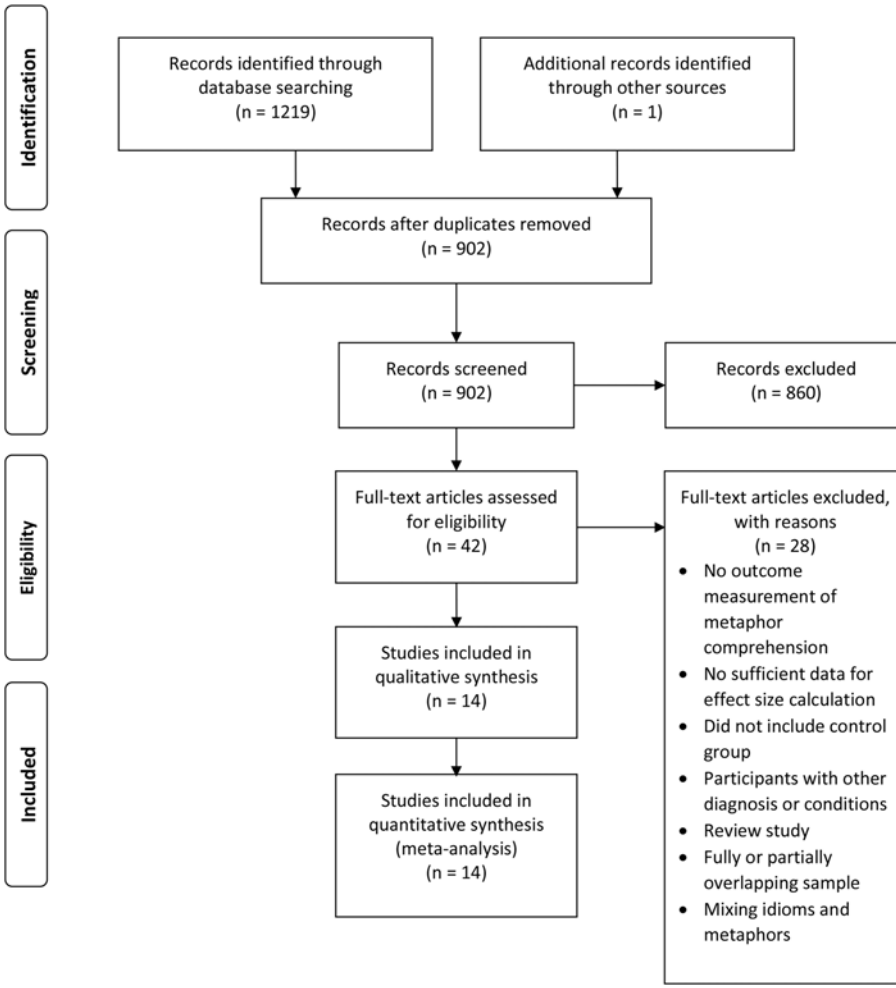

Figure 1. Flow chart for the screening and inclusion of studies based on the PRISMA statement (Moher, Liberati, Tetzlaff, & Altman, 2009).

deviations), sample sizes in each group, and means and standard deviations or *p* values for measures of metaphor comprehension. The following information about the task properties was coded: response format, metaphor familiarity, syntactic structure of the metaphor, linguistic context, and stimulus modality.

Several considerations were made when extracting the means, standard deviations, or *p* values for calculating effect sizes in the meta-analysis. First, for the studies with multiple data collection points (e.g., intervention studies), only data from the first time point was coded. This was to ensure the results were not influenced by any intervention effects. Second, to avoid estimate dependency, the data from the largest sample was extracted when overlapping samples existed (Borenstein et al., 2009). Third, to avoid the problems with assigning more weight to studies with more outcome variables (Borenstein et al., 2009), we calculated a composite score of multiple

https://doi.org/10.1017/S0142716419000328 Published online by Cambridge University Press

outcomes (e.g., novel and conventional metaphors) within each of the studies. The composite score is the mean effect size, with a variance that considers the correlation among the different outcomes. Thus, every study including multiple outcomes was represented by one score, which was used as the unit of analysis.

As predetermined, the authors initially discussed the coding procedure, then the first and the second authors double-coded the data from 10 randomly selected papers and discussed the coding of the remaining 4 papers. The interrater agreements for the coded variables in the 10 randomly selected papers were as follows: 100% for author, publication year, ASD and comparison group, age of the participants, sample size in each group, metaphor familiarity, syntactic structure of the metaphor, and linguistic context; 97% for response format and stimulus modality; 93.10% for the metaphor comprehension measures (mean with SDs and *p* values). Of note, a divergence on the metaphor comprehension measures emerged with regard to the study by Kasirer and Mashal (2014). The divergence was due to the inverted values for the ASD and TD groups reported on the table in the original article. The last author of the original paper has confirmed the typo in email correspondence. The correct values were used for calculating the effect sizes. The other disagreements between the raters were resolved by discussion and/or by consulting the original papers.

The procedure of systematically reviewing the task properties

A comprehensive coding scheme was developed for the scrutiny of the relevant data from the included studies. Data on metaphor properties were analyzed in detail for response format and linguistic characteristics (metaphor familiarity, syntactic structure, linguistic context, and stimulus modality). The exact number of studies reporting on each of these properties was identified. The findings of the studies that experimentally examined a property of interest are presented in the Results section descriptively. Lack of taking into account the properties was also considered a noteworthy finding. If the studies did not report task properties, we tried to obtain the necessary information by locating a description of the task from previous studies through searching Google web by the task name.

Meta-analytical procedure

Statistical analyses were conducted using the Comprehensive Meta-Analysis software Version 3 (Borenstein, Hedges, Higgins, & Rothstein, 2014). Because of the importance of adjusting a meta-analysis to the studies examined (Borenstein et al., 2009), we made some considerations for effect size computations. In particular, we used the Hedges' formula for standardized mean difference with a confidence interval (CI) of 95% to report effect sizes. Hedges' *g* was selected because it is corrected for sample sizes (Hedges, 1981) and studies on metaphor comprehension in ASD often include small samples. A positive Hedges' *g* value indicated that individuals with ASD had the higher group mean; a negative Hedges' *g* value indicated that the groups differed in favor of TD group. A 95% CI was calculated for each effect size to indicate whether it was significantly greater than zero. The effect is statistically significant if the CI does not cross zero. The effect sizes were

https://doi.org/10.1017/S0142716419000328 Published online by Cambridge University Press

interpreted based on Cohen's (1988) benchmarks, with effect $d \leq 0.2$ reflecting a small effect, $d \leq 0.5$ considered medium effect, and $d \leq 0.8$ indicating a large effect. However, these values are relative and somewhat arbitrary both to each other and to the specific study and research method employed (Cohen, 1988; Thompson, 2007). Therefore, interpreting these guidelines in relation to the clinical consequences that the effect size may have (Lakens, 2013) is important to avoid misleading suggestions to the practice. Hence, reporting the effect sizes in the Results section of this paper is complemented by a descriptive review.

Effect sizes across studies were averaged using a random-effects model, which does not assume that all studies in the meta-analysis share a common true effect size (Borenstein et al., 2009).

To visualize the distribution of effect sizes and CIs, and to detect possible outliers, a forest plot was used. We also performed sensitivity analysis to determine the impact of potential outliers. Sensitivity analysis makes it possible to estimate the adjusted overall effect size after removing studies one by one when extreme effect sizes are detected.

Heterogeneity

We used the Q test of homogeneity (Hedges & Olkin, 1985) to examine the heterogeneity in effect sizes. The Q statistic with its p value in a random effect model is a test of significance and reflects whether the variance is significantly different from zero. In addition, we used I^2, which reflects the extent of overlap of confidence intervals and is considered a measure of inconsistency.

Publication bias

Despite our efforts to identify gray literature, low-effect or nonsignificant studies could still be missing from the meta-analysis. To detect and statistically estimate the potential retrieval bias, we examined a funnel plot, in which a sample-size dependent statistic is plotted on the y-axis and the effect size is plotted on the x-axis. In the absence of publication bias, this plot should form a symmetrical funnel (Cooper, Hedges, & Valentine, 2009). However, the funnel plot can be difficult to interpret visually when using a random effects model (Lau, Ioannidis, Terrin, Schmid, & Olkin, 2006). Therefore, in addition, a "Trim and Fill" analysis (Duval & Tweedie, 2000) was applied. In the eventual presence of publication bias, the "Trim and Fill" analysis would be used to impute values in the funnel plot to make it symmetrical, and an adjusted overall mean effect size would be calculated.

Results

The results from the literature search are reported, followed by the description of results from the systematic review of the task properties. Finally, we present the results from the meta-analysis.

Results from the literature search

The electronic search yielded 1,219 references. In addition, one study was identified through searching in the references. All hits were screened and 14 studies

https://doi.org/10.1017/S0142716419000328 Published online by Cambridge University Press

(13 published papers and 1 conference proceeding that met the inclusion criteria were included in the systematic review and meta-analysis. Information on the screening process and the reasons for study exclusion are reported in Figure 1.

Results from the systematic review of metaphor task properties

The detailed description of the task properties of the included studies is presented in Table 2.

Response format

The answers across the tasks were elicited by the following response formats: *verbal explanation or justification*, where participants were asked to explain the meaning of the expression ($n = 2$; Melogno, D'Ardia, et al., 2012; Rundblad & Annaz, 2010a); *multiple-choice*, where participants had to choose the correct answer among a series of 3, 4, or 5 options ($n = 7$; Adachi et al., 2004; Huang, Oi, & Taguchi, 2015; Kasirer & Mashal, 2014, 2016; Mashal & Kasirer, 2011; Olofson et al., 2014; Zheng, Jia, & Liang, 2015); *meaningfulness decision*, where participants were asked to decide whether the expression makes sense or not (yes/no; $n = 4$; Chouinard & Cummine, 2016; Gold & Faust, 2010; Gunter, Ghaziuddin, & Ellis, 2002; Hermann et al., 2013). Two studies (Gunter et al., 2002; de Villiers et al., 2011) combined *multiple-choice or meaningfulness decision* and *verbal explanation/justification formats*. De Villiers et al. (2011) used multiple-choice picture modality followed by the question requiring verbal explanation. Metaphor explanation responses were reported in the results. However, the scoring strategy is not explained in their paper and, therefore, it is not clear whether the responses from the multiple-choice task have also been merged in the reported results. Gunter et al. (2002) used three tasks (multiple-choice combined with verbal explanation and meaningfulness decision task requiring to decide whether metaphors were plausible or not). However, the tasks were not described in detail in the paper, so we obtained the necessary information about the task properties by searching previous studies that employed the same tasks (Bottini et al., 1994; Jodzio, Lojek, & Bryan, 2005). Furthermore, Gunter et al. (2002) did not explain how the answers were scored and how the results obtained from the multiple-choice and verbal explanation tasks were presented in relation to each other.

None of the included studies manipulated response format in order to investigate its impact on performance.

Metaphor familiarity

Most studies employed tasks that included novel as well as conventional metaphors ($n = 7$; Gold & Faust, 2010; Gunter et al., 2002; Kasirer & Mashal, 2014, 2016; Mashal & Kasirer, 2011; Olofson et al., 2014; Zheng et al., 2015), while others included only novel ($n = 2$; Hermann et al., 2013; Melogno, D'Ardia, et al., 2012) or only conventional metaphors ($n = 1$; Rundblad & Annaz, 2010a). Four studies (Adachi et al., 2004; Chouinard & Cummine, 2016; Huang et al., 2015; de Villiers et al., 2011) did not specify metaphor familiarity.

https://doi.org/10.1017/S0142716419000328 Published online by Cambridge University Press

Table 2. Properties of the tasks employed in the included studies

Study	Response format	Linguistic characteristics 1. Metaphor familiarity 2. Syntactic structure 3. Linguistic context 4. Modality	Effect size (*g*)	95% CI	Language
Adachi et al. (2004)	Multiple-choice	1. Not specified (the words and sentences were selected from standard textbooks) 2. Mixed/not specified 3. Short stories 4. Written	−0.51	[−0.81, −0.21]	Japanese
Chouinard & Cummine (2016)	Meaningfulness decision	1. Not specified 2. Nominal 3. Sentence 4. Auditory, computer-based	−0.45	[−1.22, 0.31]	English
Gunter et al. (2002)	Multiple-choice (written) combined with verbal justification	1. Conventional 2. Not specified 3. Sentence 4. Auditory	−1.14 (combined)	[−2.17, −0.11]	English
Gunter et al. (2002) (Same study as above)	Multiple-choice (picture) combined with verbal justification	1. Conventional 2. Not specified 3. Sentence 4. Auditory	See above		English
Gunter et al. (2002) (Same study as above)	Meaningfulness decision	1. Novel (unusual) 2. Mostly nominal 3. Sentence 4. Not specified	See above		English
Gold & Faust (2010)	Meaningfulness decision	1. Novel and conventional 2. Word pairs 3. No context 4. Written, computer-based	−0.52	[−1.02, −0.01]	Hebrew
Hermann et al. (2013)	Meaningfulness decision	1. Novel 2. Nominal 3. No context 4. Written, computer-based	−0.39	[−1.00, 0.23]	German
Huang et al. (2015)	Multiple-choice	1. Not specified 2. Mixed/not specified 3. Short stories 4. Written	0.52	[−0.92, −0.13]	Taiwanese
Kasirer & Mashal (2014)	Multiple-choice	1. Novel and conventional 2. Word pairs 3. No context 4. Not specified	0.42	[−1.09, 0.26]	Hebrew

https://doi.org/10.1017/S0142716419000328 Published online by Cambridge University Press

Table 2. (*Continued*)

Study	Response format	Linguistic characteristics 1. Metaphor familiarity 2. Syntactic structure 3. Linguistic context 4. Modality	Effect size (*g*)	95% CI	Language
Kasirer & Mashal (2016)	Multiple-choice	1. Novel and conventional 2. Word pairs 3. No context 4. Not specified The same task as in Kasirer & Mashal (2014)	−0.46	[−0.92, −0.02]	Hebrew
Mashal & Kasirer (2011)	Multiple-choice	1. Novel and conventional 2. Word pairs 3. No context 4. Not specified The same task as in Kasirer & Mashal (2014) and Kasirer & Mashal (2016)	−0.76	[−1.40, −0.13]	Hebrew
Melogno, D'Ardia, et al. (2012)	Verbal explanation	1. Novel 2. Sentences 3. 12 metaphors in sentences, and 13 metaphors contextualized in four stories 4. Not specified *Junior Metaphor Comprehension Test. Designed for the specific age range and validated for use with a pediatric population*	−0.61	[−1.18, −0.43]	Italian
Olofson et al. (2014)	Multiple-choice	1. Primary conceptual novel and lexicalized 2. Sentence with metaphorical verbs or adjectives. Ex: "Now I view it!" or "Now I see it." 3. Two five-sentence stories 4. Auditory, computer-based	−0.91	[−1.69, −0.12]	English
Rundblad & Annaz (2010a)	Verbal/open question	1. Conventional 2. Sentences 3. Short picture stories 4. Auditory	−2.20	[−3.14, −1.27]	English
Zheng et al. (2015)	Multiple-choice	1. Conventional and novel 2. Nominal 3. Short stories and pictures 4. Written and illustrations	−0.71	[−1.43, 0.02]	Chinese
de Villiers et al. (2011)	Multiple-choice combined with verbal response	1. Not specified 2. Word pairs 3. Short picture stories 4. Not specified	−0.84	[−1.37, −0.31]	English

https://doi.org/10.1017/S0142716419000328 Published online by Cambridge University Press

Based on the results of the included studies, the impact of familiarity varied across studies, with some studies reporting group differences for conventional, but not for novel, metaphors (Kasirer & Mashal, 2016; Mashal & Kasirer, 2011), while others reported no group differences based on familiarity (Kasirer & Mashal, 2014). For example, some studies found that individuals with ASD could interpret both conventional metaphors (e.g., "Susan is a warm person") and novel metaphors (e.g., "Susan is a toasty person"; Olofson et al., 2014), and others found that novel metaphors were more difficult for individuals with ASD than conventional metaphors, yet this was also the case for individuals with TD (Gold & Faust, 2010; Zheng et al., 2015).

Syntactic structure

Based on those studies that provided information about syntactic structure or examples of metaphor items, the tasks varied greatly according to this variable as well. Six studies (Adachi et al., 2004; Chouinard & Cummine, 2016; Gunter et al., 2002; Hermann et al., 2013; Huang et al., 2015; Zheng et al., 2015) involved (mostly) nominal or mixed syntactic structure. Five studies (de Villiers et al., 2011; Gold & Faust, 2010; Kasirer & Mashal, 2014, 2016; Mashal & Kasirer, 2011) involved word pairs (noun–adjective pairs). Note that word-pair metaphors in de Villiers et al. (2011) were incorporated in interrogative sentence ("Which one is the *blind house*?"), while other studies did not embed word-pair metaphors in any context. Syntactic structure for conventional metaphors in Gunter et al. (2002) was not specified. Melogno, D'Ardia, et al. (2012) and Rundblad and Annaz (2010a) included sentences. Olofson et al. (2014) also included sentences with either verbs (predicate metaphors) or adjectives and was the only included study that explicitly focused on conceptual metaphors. Of note, the syntactic structure might be linked to different theoretical accounts of metaphor. For example, pragmatics-oriented scholars mostly consider "X is Y" expressions, while the literature in cognitive linguistics focuses on the multiplicity of linguistic structures that might reflect underlying conceptual metaphors and considers metaphorically used verbs or longer expressions. However, this kind of theory-driven distinction has not been considered in the literature on ASD.

Overall, because some studies failed to provide information on syntactic structure, and several papers included only a few examples of metaphors without indicating whether the metaphor task was consistent in terms of the syntactic structure, the exact number of studies using any specific syntactic structure is impossible to report. Moreover, it is important to note that there might have been inconsistent items in the data sets. For instance, Gunter et al. (2002) adopted novel (or unusual, as they are call them in the paper) metaphors from Bottini et al. (1994), which were mostly nominal (X is Y). Following our methodological choice of basing the review on what is reported by the authors in the paper, we made a judgment based on this information and classified the items used in this study as nominal. However, we are aware that at least some metaphor items are not nominal (see the metaphor examples provided by Bottini et al., 1994). Similarly, Adachi et al. (2004) used metaphors with mixed structures. Huang et al. (2015) translated the same stimuli used by

https://doi.org/10.1017/S0142716419000328 Published online by Cambridge University Press

Adachi et al. (2004) from Japanese into Taiwanese. One of the example items in both the Adachi et al. (2004) and Huang et al. (2015) studies is however translated into English as a simile. Although metaphors and similes are different figurative types and are understood differently (Happé, 1993), we decided to maintain these studies in the analysis both to be consistent with our methodological approach (basing the review on what was reported by the authors) and because the other example items in Adachi et al. (2004) were metaphors.

For all the above reasons, and also because none of the included studies manipulated syntactic structure, the impact that variation in this linguistic variable might have on the group differences in metaphor comprehension is not clear.

Linguistic context

The type of context across the studies varied from none or minimal context (word pairs or sentence-level, $n = 8$; Chouinard & Cummine, 2016; Gold & Faust, 2010; Gunter et al., 2002; Hermann et al., 2013; Kasirer & Mashal, 2014, 2016; Mashal & Kasirer, 2011; Melogno, D'Ardia, et al., 2012) to scenarios or short stories with or without accompanying pictures ($n = 6$; Adachi et al., 2004; de Villiers et al., 2011; Huang et al., 2015; Olofson et al., 2014; Rundblad & Annaz, 2010a; Zheng et al., 2015). The task employed by Melogno, D'Ardia, et al. (2012) involved metaphors presented both in decontextualized sentences and in short story context. However, no results relating to the influence of the context are reported in that study. Other studies did not manipulate the context experimentally. Thus, no results regarding the impact of linguistic context on group differences in metaphor comprehension can be reported in this review.

Stimulus modality

Five studies (Adachi et al., 2004; Gold & Faust, 2010; Hermann et al., 2013; Huang et al., 2015; Zheng et al., 2015) presented the stimuli in written modality. Four studies (Gunter et al., 2002, for the conventional metaphor task only; Chouinard & Cummine, 2016; Olofson et al., 2014; Rundblad & Annaz, 2010a) delivered metaphor comprehension aurally. Computer-based tasks were administered either aurally (Olofson et al., 2014) or in written form (Gold & Faust, 2010; Hermann et al., 2013). Five studies (Gunter et al., 2002; Kasirer & Mashal, 2014, 2016; Mashal & Kasirer, 2011; Melogno, D'Ardia, et al., 2012) did not specify the modality. Gunter et al. (2002) did not report information about the stimulus modality, but we could identify the modality (for conventional metaphors only) in the previous study (Jodzio et al., 2005). Stimulus modality is not specified in Melogno, D'Ardia, et al. (2012). De Villiers et al. (2011) employed stimuli with pictures, but without any indication whether participants were asked to read the metaphors or whether the questions were asked aurally. Three additional studies included stimulus material with pictures (Olofson et al., 2014; Rundblad & Annaz, 2010a; Zheng et al., 2015).

As a final remark, the only study that reported that they used a test validated for the age group of the participants was of Melogno, D'Ardia, et al. (2012).

https://doi.org/10.1017/S0142716419000328 Published online by Cambridge University Press

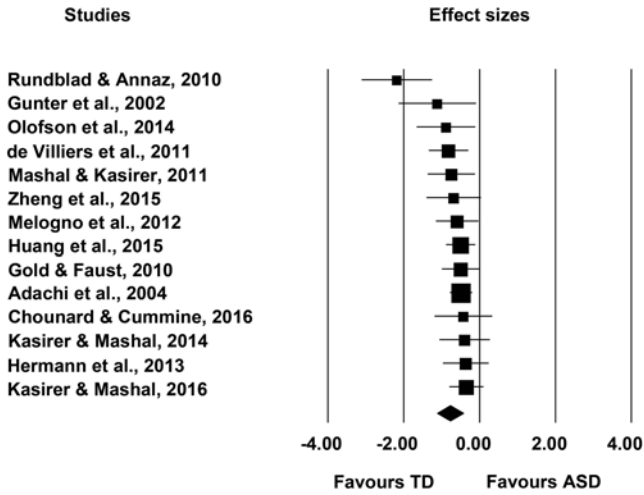

Figure 2. Forest plot showing overall average effect size for metaphor comprehension and the effect sizes with the confidence intervals for each study.

Metaphor comprehension in individuals with ASD and TD controls: A meta-analysis

Fourteen independent effect sizes, involving 336 individuals with ASD (mean sample size = 24, *SD* = 15.01, range 8–54) and 498 individuals with TD (mean sample size = 35.57, *SD* = 48.47, range 8–199), examined the differences in metaphor comprehension between the two groups. The standardized mean effect size was moderate, $g = -0.63$, 95% CI [−0.80, −0.46], $p < .001$, in favor of individuals with TD. This indicates that individuals with ASD on average have more difficulties in metaphor comprehension compared to individuals with TD.

The heterogeneity between studies was not significant, $Q\ (13) = 16.50$, $p = .22$, and 21.20% of true variability (I^2) could be explained by individual study characteristics. Higgins, Thompson, Deeks, and Altman (2003) provide some rough benchmarks for I^2, which refer to the question of what proportion of the observed variation is real. They suggest considering values below 25% as low.

Sensitivity analysis showed that the overall effect size ranged from $g = -0.66$, 95% CI [−0.85, −0.46], to $g = -0.57$, 95% CI [−0.71, −0.42]. The funnel plot showed symmetrical distribution indicating no publication bias. No studies were imputed in a Trim and Fill analysis indicating again that no publication bias was detected. The forest plot (Figure 2) shows the group differences and CIs between individuals with ASD and TD in terms of the metaphor comprehension.

Impact of response format on between-study variance

We intended to examine the response format as a potential moderator of between-study variation. However, due to the limited number of studies on each response format category (e.g., only two studies on verbal explanation format), a meta-regression or a subgroup analysis (which may be considered as a special case of meta-regression;

https://doi.org/10.1017/S0142716419000328 Published online by Cambridge University Press

Fu et al., 2011) would yield nonreliable results because of low statistical power. Specifically, it is recommended that for a categorical subgroup variable (response format in our case), each subgroup should include a minimum of four studies (Fu et al., 2011). Therefore, we qualitatively report the observed effect sizes with CIs to identify the patterns of possible relationships between response format and the heterogeneity between studies. Although not aggregated, the descriptively reported effect sizes can still guide interpretation of results and inform future studies.

Among the four types of *response format* identified in the included studies, the two studies that required verbal explanations showed moderate to large effect sizes (Melogno, D'Ardia, et al., 2012; Rundblad & Annaz, 2010a). One of these studies (Rundblad & Annaz, 2010a) generated the largest effect size from the included studies: $g = -2.20$, 95% CI [−3.14, 1.27]. This study employed an open verbal explanation task in which the short stories were accompanied with simple, hand-drawn pictures (hence, two modalities were involved). The experimenter read each story while presenting the child with one simple picture showing one story character. The child was asked to report what that character saw. In the other study that used verbal explanation response format (Melogno, D'Ardia, et al., 2012), the yielded effect size was moderate, $g = -0.62$, 95% CI [−1.18, −0.04]. This study assessed metaphor comprehension using the Junior Metaphor Comprehension Test, a validated tool for use with a pediatric population (Pinto, Melogno, & Iliceto, 2008).

Large group differences were found in the two studies that combined verbal explanation with other response formats. De Villiers et al. (2011) combined verbal justification/explanation and picture multiple-choice response formats in the same task and yielded large effect size: $g = -0.84$, 95% CI [−1.41, −0.27]. In Gunter et al. (2002) the combined effect size for the three tasks used (multiple-choice combined with verbal explanation and meaningfulness decision) was large: $g = -1.14$, 95% CI [−2.17, 0.11]. Two caveats related to this study must be mentioned: first, this study included a very small sample ($n = 8$), and second, the stimulus material in the meaningfulness task involved linguistically complex language (i.e., "The politician who didn't give straight answers was jumping ditches"; "The meaning of life is an itch you can't scratch"; or "The old man had a head full of dead leaves"; see Bottini et al., 1994, for more examples).

For the seven studies that employed multiple-choice approach only (Adachi et al., 2004; Huang et al., 2015; Kasirer & Mashal, 2014, 2016; Mashal & Kasirer, 2011; Olofson et al., 2014; Zheng et al., 2015), effect sizes varied from small to large: $g = -0.37$, 95% CI [−0.83, −0.09] to $g = -0.91$, 95% CI [−1.69, −0.12]. In the three studies that used meaningfulness decision tasks, effect sizes ranged from small to moderate: $g = -0.39$, 95% CI [−1.00, 0.23] to $g = -0.52$, 95% CI [−1.03, −0.01].

Discussion

The aim of the present systematic review and meta-analytic study was twofold: first, we sought to explore the properties of the metaphor tasks used in research involving individuals with ASD in a systematic manner. Second, we intended to examine the extent to which the groups of individuals with ASD differed from individuals with TD on metaphor comprehension, and whether any between-study variation could be explained by the properties of metaphor comprehension tasks.

https://doi.org/10.1017/S0142716419000328 Published online by Cambridge University Press

We found that the included studies employed different types of materials and tasks either invented by the researchers who designed the studies or adopted (and sometimes translated) from previous studies. Although the task properties varied greatly, the potential impact of the task properties was rarely considered. Regarding the group differences, overall, individuals with ASD fell behind their TD controls in comprehension of metaphors. The patterns show that verbal explanation response format (either pure verbal explanation or in combination with other response formats) resulted in the large effect sizes. However, due to the scarce experimental manipulation of task properties, their moderating role could not be established based on the included studies.

Properties of the tasks are seldom considered and/or controlled in the studies

In terms of the *response format*, different approaches are adopted across the studies, with the most common being multiple-choice format. Less often used response formats include verbal explanation, followed by verbal explanation combined with another response format, and meaningful decision format. It is possible that the studies involving individuals with ASD avoid using verbal explanation tasks because of known challenges related to this type of response format (i.e., cognitively, linguistically, and socially more demanding). As the impact of response format has been associated with the between-group difference in other populations (see for example, Perlini et al., 2018), we anticipated detecting similar patterns in studies comparing individuals with ASD to those with TD. However, the included studies did not experimentally manipulate the response format. Therefore, firm conclusions based on the results reported in these studies cannot be drawn.

A noteworthy finding of this review is that the impact of some of the properties, such as metaphor *familiarity*, are more frequently considered than others. The reason might be that ASD is a suitable condition for studying the distinction between novel and conventional metaphors, due to the common impairment observed in pragmatic language in this population (Paul & Norbury, 2012). Specifically, individuals with ASD should have more problems with comprehending novel as compared to conventional metaphors due to the involvement of inferential pragmatic ability to a greater extent in novel metaphors than in conventional metaphors. Yet results regarding the impact of metaphor familiarity on group difference between individuals with ASD and individuals with TD are mixed and inconclusive. This might partially be explained by different ways in which the studies have rated the degree of familiarity. For example, familiarity is often assessed based on the ratings collected from a limited number of participants, which might be not reliable given the large differences in subjective judgment on familiarity. Accordingly, Thibodeau, Sikos, and Durgin (2018) have questioned construct validity of sentence-level subjective ratings of metaphors collected from native speakers and argued that familiarity ratings are likely to be confounded with processing fluency (i.e., how easily people understand the sentences). Moreover, it may also be that other properties that covary with familiarity, such as word-level psycholinguistic characteristics (e.g., frequency, concreteness, and length), as well as metaphors' characteristics such as interpretability, naturalness, and imageability may account for distinct results (Cardillo et al., 2010). Therefore, we argue that using stimuli for which these

https://doi.org/10.1017/S0142716419000328 Published online by Cambridge University Press

properties have been rated by large samples of participants, and controlled for, could offer a more robust benchmark to explore the difference between familiar and unfamiliar metaphors. In addition, the use of controlled materials will favor the comparison of the findings across studies, which will be a great advantage for future systematic reviews and meta-analytic studies.

Syntactic structure is the least explored property and was not normally controlled for in the studies. Given that individuals with ASD frequently show impairments in syntax (Brynskov, Krøjgaard, & Eigsti, 2016), and given the evidence that metaphors in different syntactic structures are comprehended differently (Cardillo et al., 2012; Chen et al., 2008), the syntactic structure of the metaphoric items may have been important to take into account in research with individuals with ASD. However, as most studies do not report on syntactic structure (or offer inconsistent examples), we cannot conclude that the stimuli did consistently display the same structure throughout the task. Based on our results, reporting the number of studies according to the syntactic structure of the metaphors should be therefore considered with caution.

The impact of *context* on the between-group difference is also poorly explored in ASD research. This finding is striking given that inferring meaning from context has been reported to be challenging for individuals with ASD due to a cognitive difference in the normal drive for coherence (Frith, 1989; Frith & Happé, 1994; Happé, 1999; see, however, Brock, Norbury, Einav, & Nation, 2008, suggesting that differences in processing linguistic context in individuals with ASD are actually related to individual differences in their core language abilities). This implies that, although context may facilitate comprehension in TD, it may pose problems in individuals with ASD, which is in line with the "context blindness" hypothesis referring to a lack of contextual sensitivity in ASD (Vermeulen, 2014).

One study that was screened within this review (but excluded in the full-text screening stage due to the reported co-occuring conditions among individuals with ASD) found that context facilitates metaphor comprehension in ASD (Giora, Gazal, Goldstein, Fein, & Stringaris, 2012). However, one study is not enough to infer a pattern concerning the role of context. Of note, not only the presence or absence of context, but also the type of context may matter, because context with a large amount of information could hamper comprehension by overloading participants' working memory and affecting attention (see Boxhoorn et al., 2018; Pennington & Ozonoff, 1996).

Regarding *stimulus modality*, most included studies used written tasks. This could be preferable when measuring metaphor comprehension in individuals with ASD because written tasks do not pose high social interaction demands and are less taxing for memory. In addition, aurally delivered tasks might be difficult for individuals with ASD due to their characteristics in processing auditory semantic information from spoken language (see O'Connor, 2012, for a review). In contrast, it is still unclear whether written words facilitate comprehension processes for individuals with ASD in general. There is some evidence that young individuals with ASD benefit more from written word priming (not metaphorical, but conventional, "literal" words) in their lexical access than young TD controls and older individuals with ASD (Harper-Hill, Copland, & Arnott, 2014; see, however, Kamio & Toichi, 2000, suggesting the possible advantage of pictures over words in access to semantics in ASD). It is unknown if similar effects encompass the case of metaphor.

https://doi.org/10.1017/S0142716419000328 Published online by Cambridge University Press

In sum, there is a lack of attention to the role of task properties in performance on metaphor comprehension tasks in the existing ASD research. We observed the discrepancy in the task properties across the studies, as well as the limited number of studies experimentally manipulating the task properties. Therefore, strong conclusions about the extent to which task properties can explain the distinct findings in the ASD literature cannot be drawn from this study. Nevertheless, our study offers new insights into how studies in ASD have assessed metaphor comprehension and directs the focus toward the importance of acknowledging the substantial variability in tasks and their properties when interpreting the results from the existing studies. In addition, it calls for careful consideration when designing and reporting on task properties in metaphor studies.

Metaphor comprehension is more challenging for individuals with ASD than for individuals with TD

Overall, individuals with ASD as a group exhibited more difficulties in metaphor comprehension than the comparison group of individuals with TD. This finding is consistent with the results from prior studies (i.e., Happé, 1993; Rundblad & Annaz, 2010a; van Herwegen & Rundblad 2018), as well as with the findings from a recent meta-analysis (Kalandadze et al., 2018).

Taken together, this evidence indicates that, as a group, individuals with ASD more frequently experience problems in metaphor comprehension. Nevertheless, we need to acknowledge that there are several possible explanations for the significant group difference. For example, the meta-analysis (Kalandadze et al., 2018) and single studies have found that group-matching strategies could explain the between-study variation on figurative language comprehension. In particular, if ASD and TD groups were matched for language ability, the groups have been found to not differ significantly on metaphor comprehension (Norbury, 2005). These variables should necessarily be taken into account when explaining the difficulties with metaphor comprehension in individuals with ASD, together with the role of the metaphor task properties, which, despite its well-documented importance for metaphor comprehension, has not been examined until now.

Observed pattern of the associations between the response format and between-study variation

As hypothesized, verbal explanation tasks (pure verbal explanation or combined with other response formats) are, based on the observed effect sizes, most challenging for individuals with ASD as compared to TD controls. This is not surprising because explaining metaphorical meaning is cognitively, linguistically, and socially demanding, as it requires planning and formulating utterances, and thus relies on expressive language as well as metalinguistic and executive skills, which have often been found to be challenging for individuals with ASD (Hill, 2004; Kwok et al., 2015; Lewis et al., 2007). This finding also converges with results from an irony processing study (another type of figurative language) in which minimizing the verbal and pragmatic demands of the task resulted in the similar

https://doi.org/10.1017/S0142716419000328 Published online by Cambridge University Press

accuracy in judging speaker's intent for ironic criticism between the groups of individuals with ASD and TD (Pexman et al., 2011).

However, using a verbal explanation task, if validated for use with the target group, might reduce the magnitude of the group difference. For example, Melogno, D'Ardia, et al. (2012) used the test Junior Metaphor Comprehension Test designed for the specific age range (4–6 years) and validated it for use with a "pediatric population." Using the validated tool likely resulted in a smaller effect size compared to other studies that used non-validated verbal explanation tasks. The distinction between the results based on assessing metaphor comprehension with a validated versus non-validated tool is fundamental, but, unfortunately, our result is based on one only study (Melogno, D'Ardia, et al., 2012). In order to draw clearer conclusions, more studies involving validated materials to study metaphor comprehension are needed.

Meaningfulness decision tasks seem to be the least challenging for individuals with ASD when compared to TD individuals. Although somehow surprising, this might be because this type of task does not require expressive language skills, planning and formulating the responses as in verbal explanation tasks, nor inhibiting the incorrect alternatives as in multiple-choice tasks. In addition, meaningfulness decision tasks might be less socially demanding because they require less interaction with the examiner than verbal tasks. Meaningfulness decision tasks might, therefore, be less taxing for individuals with ASD than verbal explanation tasks.

Limitations

Several limitations of this study should be considered when interpreting the findings. First, the inconsistency in using a response format across the studies made it unfeasible to examine the potential moderating effect of this variable. Although descriptively reported effect sizes are informative, a meta-regression or a subgroup analysis would allow for a more accurate examination of the relationships between the response format and between-study variation. Second, none of the included studies attached the stimulus materials in appendices. Some studies presented a few examples of the metaphorical items, whereas others did not even report the examples. Although we did not have access to the full list of stimuli, the information provided in the papers was sufficient for our purposes. Future reviews that want to examine the consistency of the metaphor stimuli in the existing studies, which is definitely worth investigating, should contact the authors and request the full set of stimuli. Future reviews should also examine what types of metaphors the stimuli contain (e.g., nominal metaphors, as in "Sally is a chameleon," or conceptual metaphors, like "I see it," where "seeing" indicates "knowing" (KNOWING IS SEEING).

Another important methodological limitation was the small sample sizes in some included studies (e.g., eight participants in Gunter et al., 2002). Because larger sample sizes correspond to less sampling bias (Borenstein et al., 2009), high-powered studies would provide better effect size estimates for this meta-analysis. However, the advantage of the meta-analysis over a single study is the increased statistical power achieved via aggregating the effect sizes from multiple samples. We, therefore, propose that the magnitude of the group difference reported in

https://doi.org/10.1017/S0142716419000328 Published online by Cambridge University Press

our study gives a reliable result. These limitations will be overcome once more tightly conducted studies are available, allowing for consistent examination of potentially important variables affecting metaphor comprehension in individuals with ASD as compared to individuals with TD.

Implications

The main implication of the findings of our study for future research on metaphor comprehension is that stimuli and tasks need to be created by carefully taking into account a range of characteristics, such as the linguistic properties and response format, whose role in modifying behavioral and neural responses is well known in the literature (e.g., Bambini et al., 2014; Schmidt & Seger, 2009). Furthermore, studies should consistently examine the role of task properties experimentally to investigate their relationships with the performance among individuals with ASD and those with TD.

Of note, the different tasks may each have advantages, but at the same time they can impede comprehension if inappropriately used with individuals with ASD. For instance, multiple-choice tasks can be desirable from the psychometric perspective due to an easy and precise scoring (Rapp et al., 2018) and high reliability (see, for instance, the different reliability values of figurative language tasks-multiple choice vs. verbal explanation in Carotenuto et al., 2018). In contrast, multiple-choice tasks are more susceptible for measurement error due to the possibility of guessing the responses (Kline, 2009), as well as due to tapping more executive functions because of the need to inhibit the incorrect alternatives and select the correct one. Another aspect that should be weighed up when designing metaphor comprehension tasks in ASD research concerns the number of options provided in multiple-choice tasks. For instance, there is evidence from another pragmatic domain (i.e., scalar implicatures) that presenting two versus three options might account for the presence or absence of group differences between individuals with ASD and individuals with TD (Schaeken, Van Haeren, & Bambini, 2018).

As for verbal explanation tasks, these appear to be more sensitive than other tasks in detecting impairment in metaphor comprehension and allow us to establish with more confidence whether the metaphors were understood or not. However, it must be pointed out that verbal explanation tasks are not recommended for use with vulnerable groups because of the extra demands they pose on the participants (see Norbury, 2004). Moreover, when using verbal explanation tasks, it is important that experimenters receive adequate training in order to achieve adequate reliability in scoring responses.

In general, using metaphor tasks created ad hoc for the specific purpose of the study is often preferable for researchers, given the multifaceted nature of a metaphor. However, greater advantages would be obtained from the use of validated or/and standardized tests with good psychometric properties. Absence of tests with properties that are consistently controlled for across the studies makes comparison of the results difficult. Only one study (Melogno, D'Ardia, et al., 2012) used a validated instrument, and despite using the demanding response format of verbal explanation, the effect size yielded was smaller than in the study by Rundblad and Annaz (2010a), which used a non-validated verbal explanation task. This single observation should be investigated in future studies.

https://doi.org/10.1017/S0142716419000328 Published online by Cambridge University Press

Another important suggestion from our findings is that the stimulus materials should be attached to the published papers. In addition, providing a detailed description of the stimulus materials is essential to enable interpretation of the findings. In general, we propose that journals develop criteria for reporting metaphor studies in order to make quality appraisal of research for the readers and for future review studies possible.

Furthermore, based on the results of our study indicating limited number of studies using online methods, more high-quality studies on metaphor comprehension in ASD are needed combining offline and online comprehension methods widely used in psycholinguistic research. Considering the many demands offline tasks pose (i.e., social and linguistic), it is difficult to pinpoint the real sources of possible difficulties in metaphor comprehension when assessed offline. Online tasks such as those employing eye-tracking methodology, priming paradigms, and computerised tests can therefore add important insight to the knowledge of metaphor comprehension in ASD by measuring implicit processing (see, for example, Naigles, 2017, for innovative paradigms and methods to investigate language in ASD that could beneficially be used in metaphor research as well). For instance, priming paradigms might offer fine-grained measures of the difficulties experienced by individuals with ASD, elucidating patterns in response times (Chahboun, Vulchanov, Saldaña, Eshuis, & Vulchanova, 2017). In addition, behavioral data could profitably be combined with the data on brain functioning to explain the neurocognitive and neurolinguistic processes underlying metaphor comprehension in individuals with ASD as compared to TD controls. For instance, Gold, Faust, and Goldstein (2010) employed event-related potential recordings to examine difficulties in semantic integration in ASD. The sample in this study, however, overlapped with the sample in Gold and Faust (2010) and therefore the study was not included in the meta-analysis. We did not identify any other study with data about the brain response that met the inclusion criteria for this review.

The main practical implication of our findings is that individuals with ASD need extensive support to learn metaphor comprehension strategies explicitly and that plans on how to promote metaphor comprehension should be made. Intervention programs concerning metaphor comprehension in ASD are very few, but results are promising. For instance, Mashal and Kasirer (2011) and Melogno, Pinto, and Di Filippo (2017) used thinking maps to enhance the abstraction of semantic features in metaphors. Teachers, special educators, and speech and language therapists could capitalize on this evidence and develop strategies to stimulate metaphorical skills. To begin with, the students could be reminded that figurative language involves the use of words in nonliteral ways. Then the students could be encouraged to use their metalinguistic skills to consider the overlapping features between the topic and vehicle of the metaphor (Nippold, 2016), similarly to the approach adopted in Mashal and Kasirer (2011) and in Melogno et al. (2017). In addition, the students could be asked to collect metaphors from different sources such as advertisements and literature, including the context in which they occur (Nippold, 2016). Teachers may also incorporate metaphors of different degrees of familiarity in minimal or short story contexts and present them both aurally and in print, with and without pictures, eliciting the answers through different response formats (see Nippold, 2016, for more ideas).

Conclusions

This paper reports the systematic review and meta-analysis concerning task properties of the metaphor tasks used in ASD research and the role that they play in determining the differences between groups of individuals with ASD and those with TD in metaphor comprehension. By focusing on the impact of the task properties, this study contributes to the ongoing debate about the potential sources of between-study variation in metaphor comprehension in individuals with ASD and offers novel insights into figurative language in this population.

The included studies used an array of different tasks with a range of properties, whose impact was rarely considered and/or experimentally manipulated. Individuals with ASD in general exhibited more difficulties in metaphor comprehension than their TD counterparts, but this difference is likely to be partially related to the task properties such as the response format. Yet, more research is needed to confirm the relationship between the task properties and between-study variance.

In light of the findings of our study, we argue that future metaphor comprehension studies comparing individuals with ASD to those with TD should carefully take into account task properties such as response format and linguistic characteristics (i.e., metaphor familiarity, syntactic structure of the metaphor, linguistic context, and stimulus modality).

Consideration of task properties is also necessary in order to design appropriate educational programs to improve figurative language competence and ultimately improve communication and academic skills of individuals with ASD.

Author Contribution. Tamar Kalandadze conceptualized, designed, and administered the study; created the coding protocol; established the inclusion and exclusion criteria; collected the data; screened the titles and abstracts; read/screened full-text articles; created the coding scheme and coded the variables; analyzed and interpreted the results; drafted the manuscript; and had the main responsibility for revising and resubmitting the manuscript after peer review.

Valentina Bambini contributed to the study design, especially to the selection of the metaphor task properties and the consideration of metaphor theory aspects; did the double-coding; contributed to analysis and interpretation of the results, especially of the systematic review; contributed to the writing of the manuscript; provided feedback as well as final approval of this paper; and contributed to revising of the manuscript.

Kari-Anne B. Næss supervised the research process; contributed to the study design, especially in the meta-analysis part; contributed to the interpretation of the results; contributed to the writing of the manuscript; provided feedback as well as final approval of this paper; and contributed to revising of the manuscript.

Acknowledgments. The authors would like to thank Ingrid Lossius Falkum for her contribution in selecting the terms for our literature search. The librarians at the University of Oslo, Glenn Karlsen Bjerkenes and Magnus Heie Gregersen, deserve thanks for assisting in the literature search process. The authors would also thank Ellie Wilson at Science in English for proofreading and providing feedback on the paper.

References

References marked with an asterisk (*) indicate studies included in the systematic review and the meta-analysis.

*Adachi, T., Koeda, T., Hirabayashi, S., Maeoka, Y., Shiota, M., Wright, E. C., & Wasa, A. (2004). The metaphor and sarcasm scenario test: A new instrument to help differentiate high functioning pervasive developmental disorder from attention deficit/hyperactivity disorder. *Brain & Development*, **26**, 301–306. doi: 10.1016/S0387-7604(03)00170-0

American Psychiatric Association. (1994). *Diagnostic and statistical manual of mental disorders: DSM-IV*. Washington, DC: Author.

American Psychiatric Association. (2000). *Diagnostic and statistical manual of mental disorders: DSM-IV-TR*. Washington, DC: Author.

American Psychiatric Association. (2013). *Diagnostic and statistical manual of mental disorders: DSM-V*. Arlington, VA: Author.

Arcara, G., Tonini, E., Muriago, G., Mondin, E., Sgarabottolo, E., Bertagnoni, G., ... Bambini, V. (2019). Pragmatics and figurative language in individuals with traumatic brain injury: Fine-grained assessment and relevance-theoretic considerations. *Aphasiology*. Advance online publication. doi: 10.1080/02687038.2019.1615033

Bambini, V., Bertini, C., Schaeken, W., Stella, A., & Di Russo, F. (2016). Disentangling metaphor from context: An ERP study. *Frontiers in Psychology*, **7**, 559. doi: 10.3389/fpsyg.2016.00559

Bambini, V., Gentili, C., Ricciardi, E., Bertinetto, P. M., & Pietrini, P. (2011). Decomposing metaphor processing at the cognitive and neural level through functional magnetic resonance imaging. *Brain Research Bulletin*, **86**, 203–216. doi: 10.1016/j.brainresbull.2011.07.015

Bambini, V., Ghio, M., Moro, A., & Schumacher, P. B. (2013). Differentiating among pragmatic uses of words through timed sensicality judgments. *Frontiers in Psychology*, **4**, 938. doi: 10.3389/fpsyg.2013.00938

Bambini, V., Resta, D., & Grimaldi, M. (2014). A dataset of metaphors from the Italian literature: Exploring psycholinguistic variables and the role of context. *PLOS ONE*, **9**, e105634. doi: 10.1371/journal.pone.0105634

Bishop, D. V. M., & Norbury, C. F. (2005). Executive functions in children with communication impairments, in relation to autistic symptomatology: 2. Response inhibition. *Autism*, **9**, 29–43. doi: 10.1177/1362361305049028

Blasko, D. G., & Connine, C. M. (1993). Effects of familiarity and aptness on metaphor processing. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, **19**, 295–308. doi: 10.1037/0278-7393.19.2.295

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

Borenstein, M., Hedges, L., Higgins, J., & Rothstein, H. (2014). *Comprehensive meta-analysis. Version 3*. Englewood, NJ: Biostat.

Bottini, G., Corcoran, R., Sterzi, R., Paulesu, E., Schenone, P., Scarpa, P., ... Frith, C. D. (1994). The role of the right hemisphere in the interpretation of figurative aspects of language: A positron emission tomography activation study. *Brain*, **117**, 1241–1253. doi: 10.1093/brain/117.6.1241

Bowdle, B. F., & Gentner, D. (2005). The career of metaphor. *Psychological Review*, **112**, 193–216. doi: 10.1037/0033-295X.112.1.193

Boxhoorn, S., Lopez, E., Schmidt, C., Schulze, D., Hänig, S., & Freitag, C. M. (2018). Attention profiles in autism spectrum disorder and subtypes of attention-deficit/hyperactivity disorder. *European Child & Adolescent Psychiatry*, **27**, 1433–1447. doi: 10.1007/s00787-018-1138-8

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. doi: 10.1016/j.cognition.2008.06.007

Brynskov, C., Krøjgaard, P., & Eigsti, I. (2016). Language and communication in children with autism: Do research and clinical practice converge in Denmark? *Nordic Psychology*, **68**, 100–113. doi: 10.1080/19012276.2015.1071199

Camp, E. (2009). Metaphor. In L. Cummings (Ed.), *The pragmatics encyclopedia*. Oxford: Routledge.

Cardillo, E. R., Schmidt, G. L., Kranjec, A., & Chatterjee, A. (2010). Stimulus design is an obstacle course: 560 matched literal and metaphorical sentences for testing neural hypotheses about metaphor. *Behavioral Research Methods*, **42**, 651–664. doi: 10.3758/BRM.42.3.651

Cardillo, E. R., Watson, C. E., & Chatterjee, A. (2017). Stimulus needs are a moving target: 240 additional matched literal and metaphorical sentences for testing neural hypotheses about metaphor. *Behavioral Research*, **49**, 471–483. doi: 10.3758/s13428-016-0717-1

Cardillo, E. R., Watson, C. E., Schmidt, G. L., Kranjec, A., & Chatterjee, A. (2012). From novel to familiar: Tuning the brain for metaphors. *Neuroimage*, **59**, 3212–3221. doi: 10.1016/j.neuroimage.2011.11.079

Carotenuto, A., Arcara, G., Orefice, G., Cerillo, I., Giannino, V., Rasulo, M., ... Bambini, V. (2018). Communication in multiple sclerosis: Pragmatic deficit and its relation with cognition and social cognition. *Archives of Clinical Neuropsychology*, **33**, 194–205. doi: 10.1093/arclin/acx061

Carston, R. (2010). Lexical pragmatics, ad hoc concepts and metaphor: A relevance theory perspective. *Italian Journal of Linguistics*, **22**, 153–180.

Carston, R. (2012). Metaphor and the literal/nonliteral distinction. In K. Allan & K. Jaszczolt (Eds.), *Cambridge handbook of pragmatics* (pp. 469–492). Cambridge: Cambridge University Press.

Chahboun, S., Vulchanov, V., Saldaña, D., Eshuis, H., & Vulchanova, M. (2017). Can you tell it by the prime? A study of metaphorical priming in high-functioning autism in comparison with matched controls. *International Journal of Language & Communication Disorders*, **52**, 766–785. doi: 10.1111/1460-6984.12314

Chakrabarty, M., Sarkar, S., Chatterjee, A., Ghosal, M., Guha, P., & Deogaonkar, M. (2014). Metaphor comprehension deficit in schizophrenia with reference to the hypothesis of abnormal lateralization and right hemisphere dysfunction. *Language Sciences*, **44**, 1–14. doi: 10.1016/j.langsci.2014.01.002

Chen, E., Widick, P., & Chatterjee, A. (2008). Functional anatomical organization of predicate metaphor processing. *Brain and Language*, **107**, 194–202. doi: 10.1016/j.bandl.2008.06.007

*__Chouinard, B., & Cummine, J.__ (2016). All the world's a stage: Evaluation of two stages of metaphor comprehension in people with autism spectrum disorder. *Research in Autism Spectrum Disorders*, **23**, 107–121. doi: 10.1016/j.rasd.2015.12.008

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences*. New York: Routledge Academic.

Cooper, H. M., Hedges, L. V., & Valentine, J. C. (Eds.) (2009). *The handbook of research synthesis and meta-analysis* (2nd ed.). New York: Sage.

Dettmer, S., Simpson, R. L., Smith Myles, B., & Granz, J. B. (2000). The use of visual supports to facilitate transitions of students with autism. *Focus on Autism and Other Developmental Disabilities*, **15**, 163–169. doi: 10.1177/1088357614528799

*__de Villiers, P. A., de Villiers, J. G., Diaz, S., Cheung, C., Alig, R., & Raditz, V.__ (2011). *Nonliteral language and theory of mind in autism spectrum disorders*. Poster presented at the American Speech-Language-Hearing Association convention, San Diego, CA, November 17–19.

Duval, S., & Tweedie, R. (2000). Trim and fill: A simple funnel-plot-based method of testing and adjusting for publication bias in meta-analysis. *Biometrics*, **56**, 455–463. doi: 10.1111/j.0006-34IX.2000.00455.x

Frith, U. (1989). A new look at language and communication in autism. *International Journal of Language and Communication Disorders*, **24**, 123–150. doi: 10.3109/13682828909011952

Frith, U., & Happé, F. (1994). Autism: Beyond "theory of mind." *Cognition*, **50**, 115–132. doi: 10.1016/0010-0277(94)90024-8

Fu, R., Gartlehner, G., Grant, M., Shamliyan, T., Sedrakyan, A., Wilt, T. J., … **Trikalinos, T. A.** (2011). Conducting quantitative synthesis when comparing medical interventions: AHRQ and the Effective Health Care Program. *Journal of Clinical Epidemiology*, **64**, 1187–1197. doi: 10.1016/j.jclinepi.2010.08.010

Gentner, D., Bowdle, B. F., Wolff, P., & Boronat, C. (2001). Metaphor is like analogy. In D. Centner, K. J. Holyoak, & B. N. Kokinov (Eds.), *The analogical mind: Perspectives from cognitive science* (pp. 199–253). Cambridge MA, MIT 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.

Gibbs, R. W. (2011). Evaluating conceptual metaphor theory. *Discourse Processes*, **48**, 529–562. doi: 10.1080/0163853X.2011.606103

Gibbs, R. W., & Tendahl, M. (2006). Cognitive effort and effects in metaphor comprehension: Relevance theory and psycholinguistics. *Mind & Language*, **21**, 379–403. doi: 10.1111/j.1468-0017.2006.00283.x

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**, 22–54. doi: 10.1080/10926488.2012.638823

Gluksberg, S. (2001). *Understanding figurative language: From metaphors to idioms*. Oxford: Oxford University Press.

Gluksberg, S. (2003). The psycholinguistics of metaphor. *Trends in Cognitive Sciences*, **7**, 92–96. doi: 10.1016/S1364-6613(02)00040-2

Gluksberg, 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. doi: 10.1016/S0022-5371(82)90467-4

Glucksberg, S., & Keysar, B. (1990). Understanding metaphorical comparisons: Beyond similarity. *Psychological Review*, **97**, 3–18. doi: 10.1037/0033-295X.97.1.3

*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. doi: 10.1007/s10803-009-0930-1

Gold, R., Faust, M., & Goldstein, A. (2010). Semantic integration during metaphor comprehension in Asperger syndrome. *Brain and Language*, **113**, 124–134. doi: 10.1016/j.bandl.2010.03.002

Golden, A. (2010). Grasping the point. A study of 15-year-old students' comprehension of metaphorical expression in schoolbooks. In G. Low, Z. Todd, A. Deignan, & L. Cameron (Eds.), *Researching and applying metaphor in the real world*. Amsterdam: Benjamins.

*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. doi: 10.1023/A:1016326701439

Happé, F. G. E. (1993). Communicative competence and theory of mind in autism: A test of relevance theory. *Cognition*, **48**, 101–119. doi: 10.1016/0010-0277(93)90026-R

Happé, F. (1999). Autism: Cognitive deficit or cognitive style? *Trends in Cognitive Sciences*, **3**, 216–222. doi: 10.1016/S1364-6613(99)01318-2

Harper-Hill, K., Copland, D., & Arnott, W. (2014). Pathways to meaning: Written and spoken word priming in children with ASD versus typically developing peers. *Research in Autism Spectrum Disorders*, **8**, 1351–1363. doi: 10.1016/j.easd.2014.07.004

Hedges, L. V. (1981). Distribution theory for Glass's estimator of effect size and related estimators. *Journal of Educational and Behavioral Statistics*, **6**, 107–128. doi: 10.3102%2F10769986006002107

Hedges, L., & Olkin, I. (1985). *Statistical methods for meta-analysis*. San Diego, CA: Academic Press.

*Hermann, I., Haser, V., van Elst, L. T., Ebert, D., Müller-Feldmeth, D., Riedel, A., & Konieczny, L. (2013). Automatic metaphor processing in adults with Asperger syndrome: A metaphor interference effect task. *European Archives of Psychiatry and Clinical Neurosciences*, **263** (Suppl. 2), S177–S187. doi: 10.1007/s00406-013-0453-9

Higgins, J. P. T., Thompson, S. G., Deeks, J. J., & Altman, D. G. (2003). Measuring inconsistency in meta-analyses. *BMJ*, **327**, 557–560.

Hill, E. L. (2004). Evaluating the theory of executive dysfunction in autism. *Developmental Review*, **24**, 189–233. doi: 10.1016/j.dr.2004.01.001

*Huang, S., Oi, M., & Taguchi, A. (2015). Comprehension of figurative language in Taiwanese children with autism: The role of theory of mind and receptive vocabulary. *Clinical Linguistics and Phonetics*, **29**, 764–775. doi: 10.3109/02699206.2015.1027833

Jacobs, A. M., & Kinder, A. (2017). "The brain is the prisoner of thought": A machine-learning assisted quantitative narrative analysis of literary metaphors for use in neurocognitive poetics. *Metaphor and Symbol*, **32**, 139–160. doi: 10.1080/10926488.2017.1338015

Jodzio, K., Lojek, E., & Bryan, K. (2005). Functional and neuroanatomical analysis of extralinguistic disorders in right hemisphere-damaged patients. *Psychology of Language and Communication*, **9**, 55–73. doi: 10.3389/fpsyg.2013.00863

Kalandadze, T., Norbury, C., Nærland, T., & Næss, B. K.-A. (2018). Figurative language comprehension in individuals with autism spectrum disorder: A meta-analytic review. *Autism*, **22**, 99–117. doi: 10.1177/1362361316668652

Kamio, Y., & Toichi, M. (2000). Dual access to semantics in autism: Is pictorial access superior to verbal access? *Journal of Child Psychology and Psychiatry, and Allied Disciplines*, **41**, 859–867. doi: 10.1017/S0021963099006137

*Kasirer, A., & Mashal, N. (2014). Verbal creativity in autism: comprehension and generation of metaphoric language in high-functioning autism spectrum disorder and typical development. *Frontiers in Human Neuroscience*, **8**, 615. doi: 10.3389/fnhum.2014.00615

*Kasirer, A., & Mashal, N. (2016). Comprehension and generation of metaphors by children with autism spectrum disorder. *Research in Autism Spectrum Disorders*, **32**, 53–63. doi: 10.1016/j.rasd.2016.08.003

Kline, A. B. (2009). *Becoming a behavioural science researcher. A guide to producing research that matters*. New York: Guilford Press.

Kwok, E. Y. L., Brown, H. M., Smyth, R. E., & Cardy, J. O. (2015). Meta-analysis of receptive and expressive language skills in autism spectrum disorder. *Research in Autism Spectrum Disorders*, **9**, 202–222. doi: 10.1016/j.rasd.2014.10.008

Lakens, D. (2013). Calculating and reporting effect sizes to facilitate cumulative science: A practical primer for t-tests and ANOVAs. *Frontiers in Psychology*, **4**, 863.

Lakoff, G., & Johnson, M. (1980). *Metaphors we live by: Language*. Chicago: University of Chicago Press.

Lau, J., Ioannidis, J., Terrin, N., Schmid, C. H., & Olkin, I. (2006). The case of the misleading funnel plot. *British Medical Journal*, **333**, 597–600. doi: 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. doi: 10.1017/S030500091800048X

Lee, S. S., & Dapretto, M. (2006). Metaphorical vs. literal word meanings: FMRI evidence against a selective role of the right hemisphere. *NeuroImage*, **29**, 536–544. doi: 10.1016/j.neuroimage.2005.08.003

Lewis, F. M., Murdoch, B. E., & Woodyatt, G. C. (2007). Communicative competence and metalinguistic ability: Performance by children and adults with autism spectrum disorder. *Journal of Autism and Developmental Disorders*, **37**, 1525–1538. doi: 10.1007/s10803-006-0265-0

Mashal, N., Faust, M., Hendler, T., & Jung-Beeman, M. (2009). An fMRI study of processing novel metaphoric sentences. *Laterality*, **14**, 30–54. doi: 10.1080/13576500802049433

*__Mashal, N., & Kasirer, A.__ (2011). Thinking maps enhance metaphoric competence in children with autism and learning disabilities. *Research in Developmental Disabilities*, **32**, 2045–2054. doi: 10.1016/j.ridd.2011.08.012

*__Melogno, S., D'Ardia, C., Pinto, M. A., & Levi, G.__ (2012). Explaining metaphors in high-functioning autism spectrum disorder children: A brief report. *Research in Autism Spectrum Disorders*, **6**, 683–689. doi: 10.1016/j.rasd.2011.09.005

Melogno, S., Pinto, M. A., & Di Filippo, G. (2017). Sensory and physico-psychological metaphor comprehension in children with ASD: A preliminary study on the outcomes of a treatment. *Brain Sciences*, **7**, 85. doi: 10.3390/brainsci7070085

Melogno, S., Pinto, M. A., & Levi, G. (2012). Metaphor and metonymy in ASD children: A critical review from a developmental perspective. *Research in Autism Spectrum Disorders*, **6**, 1289–1296. doi: 10.1016/j.rasd.2012.04.004

Melogno, S., Pinto, M. A., & Levi, G. (2015). Profile of the linguistic and metalinguistic abilities of a gifted child with autism spectrum disorder: A case study. *Child Language Teaching and Therapy*, **31**, 113–126. doi: 10.1177/0265659014530414clt

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 Med*, **6**, e1000097. doi: 10.1371/journal.pmed1000097

Naigles, L. R. (Ed.) (2017). *Innovative investigations of language in autism spectrum disorder*. Washington, DC: American Psychological Association.

Nelson, C., McDonnell, A. P., Johnston, S. S., Crompton, A., & Nelson, A. R. (2007). Keys to play: A strategy to increase the social interactions of young children with autism and their typically developing peers. *Education and Training in Developmental Disabilities*, **42**, 165–181.

Nippold, M. A. (2016). *Later language development: School-age children, adolescents, and young adults* (4th ed.). Austin, TX: Pro-Ed.

Norbury, C. F. (2004). Factors supporting idiom comprehension in children with communication disorders. *Journal of Speech, Language, and Hearing Research*, **47**, 1179–1193.

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. doi: 10.1348/026151005X26732

O'Connor, K. (2012). Auditory processing in autism spectrum disorder: A review. *Neuroscience and Biobehavioral Reviews*, **36**, 836–854. doi: 10.1016/j.neubiorev.2011.11.008

*__Olofson, E. L., Casey D., Oluyedun, O. A., van Herwegen, J., Becerra, A., & Rundblad, G.__ (2014). Youth with autism spectrum disorder comprehend lexicalized and novel primary conceptual metaphors. *Journal of Autism and Developmental Disorders*, **44**, 2568–2583. doi: 10.1007/s10803-014-2129-3

Paul, R., & Norbury, C. (2012). *Language disorders from infancy through adolescence: Listening, speaking, reading, writing, and communicating*. St. Louis, MO: Elsevier Health Sciences.

Pennington, B. F., & Ozonoff, S. (1996). Executive functions and developmental psychopathology. *Journal of Child Psychology and Psychiatry*, **37**, 51–87. doi: 10/1111/j.1469-7610.1996.tb01380.x

Perlini, C., Bellani, M., Finos, L., Lasalvia, A., Bonetto, C., Scocco, P., … GET UP group. (2018). Non-literal language comprehension in a large sample of first episode psychosis patients in adulthood. *Psychiatry Research*, **260**, 78–89. doi: 10.1016/j.psychres.2017.11.032

Pexman, P. M., Rostad, K. R., McMorris, C. A., Climie, E. A., Stowkowy, J., & Glenwright, M. R. (2011). Processing of ironic language in children with high-functioning autism spectrum disorder. *Journal of Autism and Developmental Disorders*, **41**, 1097–1112. doi: 10.1007/s10803-010-1131-7

Pinto, M. A., Melogno, S., & Iliceto, P. (2008). *TCM Junior: Test di comprensione delle metafore (Scuola dell'infanzia e scuola primaria)*. Roma: Carocci-Faber.

Pouscoulous, N. (2011). Metaphor: For adults only? *Belgian Journal of Linguistics*, **25**, 51–79. doi: 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*. Amsterdam: Benjamins.

Rao, S. M., & Gagie, B. (2006). Learning through seeing and doing. Visual supports for children with autism. *Teaching Exceptional Children*, **38**, 26–33.

Rapp, A. M., Felsenheimer, A. K., Langohr, K., & Klupp, M. (2018). The comprehension of familiar and novel metaphoric meanings in schizophrenia: A pilot study. *Frontiers in Psychology*, **8**, 2251. doi: 10.3389/fpsyg.2017.02251

Recanati, F. (2004). *Literal meaning*. Cambridge: Cambridge University Press.

Reed, P., & McCarthy, J. (2011). Cross-modal attention-switching is impaired in autism spectrum disorders. *Journal of Autism and Developmental Disorders*, **42**, 947–953. doi: 10.1007/s10803-011-1324-8

Rinaldi, M. C., Marangolo, P., & Baldassarri, F. (2004). Metaphor comprehension in right brain-damaged patients with visuo-verbal and verbal material: A dissociation (re)considered. *Cortex*, **40**, 479–490. doi: 10.1016/S0010-9452(08)70141-2

Rossetti, I., Brambilla, P., & Papagno, C. (2018). Metaphor comprehension in schizophrenic patients. *Frontiers in Psychology*, **9**, 670. doi: 10.3389/fpsyg.2018.00670

*__Rundblad, G., & Annaz, D.__ (2010a). The atypical development of metaphor and metonymy comprehension in children with autism. *Autism*, **14**, 29–46. doi: 10.1177/1362361309340667

Rundblad, G., & Annaz, D. (2010b). Development of metaphor and metonymy comprehension: Receptive vocabulary and conceptual knowledge. *British Journal of Developmental Psychology*, **28**, 547–563. doi: 10.1348/026151009X454373

Schaeken, W., Van Haeren, M., & Bambini, V. (2018). The understanding of scalar implicatures in children with Autism Spectrum Disorder: Dichotomized responses to violations of informativeness. *Frontiers in Psychology*, **9**, 1266. doi: 10.3389/fpsyg.2018.01266

Schmidt, W. L., & Seger, C. A. (2009). Neural correlates of metaphor processing: The roles of figurativeness, familiarity and difficulty. *Brain and Cognition*, **7**, 375–386. doi: 10.1016/j.bandc.2009.06.001

Siqueira, M., Marques, D. F., & Gibbs, R. W., Jr. (2016). Metaphor-related figurative language comprehension in clinical populations: A critical review. *SCRIPTA, Belo Horizonte*, **20**, 36–60.

Sperber, D., & Wilson, D. (2012). A deflationary account of metaphors. In D. Wilson & D. Sperber (Eds.), *Meaning and relevance* (pp. 97–122). Cambridge: Cambridge University Press.

Steen, G., Dorst, A., & Hermann, J. (2010). Metaphor in usage. *Cognitive Linguistics*, **21**, 765–796. doi: 10.1515/cogl.2010.024

Tager-Flusberg, H. (1992). Autistic children's talk about psychological states: Deficits in the early acquisition of a theory of mind. *Child Development*, **63**, 161–172. doi: 10.1111/j.1467-8624.1992.tb03604.x

Tager-Flusberg, H., Calkings, S., Nolin, T., Baumberger, T., Anderson, M., & Chadwick-Dias, D. (1990). A longitudinal study of language acquisition in autistic and Down's syndrome children. *Journal of Autism and Developmental Disorders*, **20**, 1–21. doi: 10.1007/BF02206853

Thibodeau, P. H., Sikos, L., & Durgin, F. H. (2018). Are subjective ratings of metaphors a red herring? The big two dimensions of metaphoric sentences. *Behavior Research Methods*, **50**, 759–772. doi: 10.3758/s13428-017-0903-9

Thompson, B. (2007). Effect sizes, confidence intervals, and confidence intervals for effect sizes. *Psychology in the Schools*, **44**, 423–432. doi: 10.1002/pits.20234

Utsumi, A., & Sakamoto, M. (2007). *Computational evidence for two-stage categorization as a process of adjective metaphor comprehension*. Paper presented at the 2nd European Cognitive Science Conference, Delphi, Greece, May 23–27.

Utsumi, A., & Sakamoto, M. (2011). Indirect categorization as a process of predicative metaphor comprehension. *Metaphor and Symbol*, **26**, 299–313. doi: 10.1080/10926488.2011.609120

van Herwegen, J., & Rundblad, G. (2018). A cross-sectional and longitudinal study of novel metaphor and metonymy comprehension in children, adolescents, and adults with autism spectrum disorder. *Frontiers in Psychology*, **9**, 945. doi: 10.3389/fpsyg.2018.00945

Varga, E., Schnell, Z., Tényi, T., Németh, N., Simon, M., Hajnal, A., ... Herold, R. (2014). Compensatory effect of general cognitive skills on non-literal language processing in schizophrenia: A preliminary study. *Journal of Neurolinguistics*, **29**, 1–16. doi: 10.1016/j.jneuroling.2014.01.001

Vermeulen, P. (2014). Context blindness in autism spectrum disorder: Not using the forest to see the trees as trees. *Focus on Autism and Other Developmental Disabilities*, **30**, 1–11. doi: 10.1177/1088357614528799

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. doi: 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*. Advance online publication. doi: 10.1007/s10849-019-09282-7

Vulchanova, M., Saldaña, D., Chahboun, S., & Vulchanov, V. (2015). Figurative language processing in atypical populations: The ASD perspective. *Frontiers in Human Neuroscience*, **9**, 24. doi: 10.3389/fnhum.2015.00024

Wilson, D. (2011). Parallels and differences in the treatment of metaphor in relevance theory and cognitive linguistics. *Intercultural Pragmatics*, **8**, 177–196. doi: 10.2478/v10148-011-0025-1

Wilson, D., & Carston, R. (2006). Metaphor, relevance and the "emergent Property." *Mind & Language*, **21**, 404–433. doi: 10.1111/j.1468-0017.2006.00284

Winner, E. (1988). *The point of words: Children's understanding of metaphor and irony*. Cambridge, MA: Harvard University Press.

Winner, E., Rosenstiel, A. K., & Gardner, H. (1976). The development of metaphoric understanding. *Developmental Psychology*, **12**, 289–297.

World Health Organization. (1992). *The ICD-10 classification of mental and behavioural disorders: Clinical descriptions and diagnostic guidelines*. Geneva: World Health Organization. http://www.who.int/iris/handle/10665/37958

*__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. doi: 10.1016/j.rasd.2014.11.007

Cite this article: 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