Abstract
Metaphor comprehension was studied in three groups of children from 6-12 years old: a group with autism spectrum disorder with level 2 severity (ASD, n=22) and two comparison groups with typical development: one matched with the ASD group on chronological age (TCD group, n=22) and the other matched on linguistic age (TLD group, n=22). The TCD group performed better than the TLD group, which performed better than the ASD group, on the comprehension of both conventional and novel metaphors, with better performance found on conventional metaphors than on novel ones. We suggest that both linguistic and extralinguistic competencies (usually limited in level 2 ASD) would be necessary for understanding metaphors, which would be facilitated by their frequency and familiarity.

Highlights:
Metaphor comprehension was worse in children with ASD than in younger TD children matched on LA.
Metaphor comprehension was worse in children with ASD than in TD children matched on CA.
The comprehension of conventional metaphors was better than the comprehension of novels.
Autism spectrum disorder (ASD) is a neurodevelopmental disorder characterized by impairments in social interaction and communication across multiple contexts and by the presence of repetitive and stereotyped behaviors (DSM-5; American Psychiatric Association -APA-, 2013). The severity of these two criteria determines the severity of the disorder, with three possible levels, 1, 2, and 3: level 1 would be the least severe (the person needs support), and level 3 would be the most severe (the person needs considerable support). When diagnosing ASD, it is necessary to determine whether there is another disorder or associated condition, such as a language impairment. Although some individuals with autism have preserved structural language skills (morphology, syntax, and phonology), pragmatic and non-literal language skills are generally affected across the autism spectrum (Baixauli-Fortea, Miranda-Casas, Berenguer-Forner, Colomer-Diago, & Roselló-Miranda, 2019; Gold & Faust, 2010; Ramberg, Ehlers, Nydén, Johansson & Gillberg, 2011).

The comprehension of non-literal (or non-explicit) language -including figurative language- requires the ability to go beyond what is explicitly stated. Difficulties in this ability can interfere with the natural flow of social interactions and impair communication (Swineford, Thurm, Baird, Wetherby & Swedo, 2014), thus constituting a risk factor for the occurrence of certain situations, such as victimization, especially in children and adolescents (Paul, 2015). Therefore, it is particularly important to study this type of language in vulnerable populations, such as individuals with ASD, whose difficulties with figurative language have been documented (Kalandadze, Norbury, Nærland & Næss, 2018; Vulchanova, Saldaña, Chahboun & Vulchanov, 2015).

Metaphors are some of the most common examples of figurative language. In this paper, we will follow the proposals of Lakoff & Johnson (2003), corroborated by Gibbs, Beitel, Harrington & Sanders (1994), Gibbs & Colston (1995), Mandler (1988, 1995, 2005, 2010), or Richardson, Spivey, Barsalou & McRae (2003). For these authors, the metaphor is a cognitive resource that involves linking two apparently unrelated knowledge domains (the base term and the target term), functioning as an everyday tool to explain one domain in terms of another. The metaphor makes it possible to organize experience based on familiar concepts through a relationship of similarity between the domains involved. The metaphor is not, therefore, a poetic resource, but rather a way of categorizing reality, and it can be present in very frequent expressions.

**Metaphor comprehension in children with and without ASD**
The comprehension of figurative language, and specifically metaphors, is a gradual process that, in children with typical development (TD), begins in early childhood, between three and four years old, and continues throughout adolescence, culminating in adulthood (Cacciari, 2014; Özçalışkan, 2005; Reyna & Kiernan, 1995; Rundblad & Annaz, 2010a; Winner, 1997). According to Pouscoulous (2011), the development of linguistic skills, the accumulation of knowledge and experiences about the world, and the development of mentalist skills and social comprehension, all of which take place with age, occur in parallel with better comprehension of figurative language and, particularly, metaphoric language.

However, children with ASD usually tend to make literal interpretations of statements, presenting difficulties in the comprehension of figurative language and, specifically, metaphoric language, as revealed in the meta-analytic review carried out by Kalandadze et al., (2018). The 41 articles included in this study compared figurative language in subjects with and without ASD, with metaphors specifically addressed in 15 of the articles. The majority of the studies that compared metaphor comprehension in individuals with and without ASD used subjects between 6 and 12 years old, as in the present study. In all of them, better performance on metaphor comprehension was obtained in children with TD than in children with ASD (Adachi et al., 2004; De Villiers et al., 2011; Dennis, Lazenby & Lockyer, 2001; Huang, Oi & Taguchi, 2015; Landa & Goldberg, 2005; Mashal & Kasirer, 2011; Olofson et al., 2014; Rundblad & Annaz, 2010b; Zheng, Jia & Liang, 2015). All these studies used the equating strategy of matching the groups on chronological age (CA). Moreover, in four studies (Adachi et al., 2004; De Villiers et al., 2011; Landa & Goldberg, 2005; Zheng et al., 2015), the groups were also matched on verbal IQ, in one study (Huang et al., 2015), they were matched on vocabulary, and in only one study (Mashal & Kasirer, 2011), they were also matched on linguistic age (LA).

In this meta-analytic review (Kalandadze et al., 2018), the largest between-group differences and the largest mean effect sizes were obtained in the studies where the groups were matched on CA, but not on indicators of linguistic ability (Dennis et al., 2001; Olofson et al., 2014; Rundblad & Annaz, 2010b). However, when the groups were matched on indicators of linguistic skills, in addition to CA, the differences between groups and the effect sizes were smaller. These results provide evidence that figurative language comprehension deficits could be related to the individual’s structural (grammar and semantic) language skills, such as vocabulary and syntax (Gernsbacher & Pripas-

However, in the studies included in the meta-analysis by Kalandadze et al. (2018) that matched the groups on CA and indicators of linguistic skills, although the between-group differences and mean effect sizes were small, they continued to be statistically significant. Likewise, in other more recent studies—which included other age groups— an atypical (delayed) processing of metaphoric language was obtained in people with level 1 ASD with intact formal linguistic skills (Chahboun, Vulchanov, Saldaña, Eshuis, & Vulchanova, 2016, 2017; Chouinard & Cummine, 2016). Therefore, in addition to formal linguistic competence, other cognitive skills (extralinguistic) would be involved in metaphor comprehension, such as mind reading, social comprehension, executive functioning (Landa & Goldberg 2005), and context processing (Vulchanova et al., 2015), skills that can be limited or compromised in ASD (Baixauli-Fortea et al., 2019; Frith & Happé, 1994; Happé & Frith, 2006; Hill, 2004; Ozonoff, 1997). In fact, in the initial studies on metaphors in ASD, the theory of mind (the capacity to interpret the speaker's mental states) was a prominent factor (Happé, 1993, 1995). Other proposals were weak central coherence (a limited ability to attend to the overall form or meaning, understand context, or "see the big picture") and executive dysfunction (deficits in a set of cognitive abilities involved in goal formulation, planning, organization, sequencing, and self-regulation) as possible factors involved in the difficulties of children with ASD in understanding metaphors (Huang et al., 2015; Kasirer & Mashal, 2014; Mashal & Kasirer, 2011; Rundblad & Annaz, 2010b).

With regard to the developmental course, age is a variable that does not seem to be significantly associated with an improvement in metaphor comprehension in children with ASD, unlike in children with TD (Huang et al., 2015; Kalandadze et al., 2018; Olofson et al., 2014; Rundblad & Annaz, 2010b; Van Herwegen & Rundblad, 2018; Whyte & Nelson, 2015). Thus, individuals with ASD did not show the same evolution in metaphor comprehension throughout childhood and adolescence as individuals with TD, with the former remaining at floor performance into adolescence, which has been labeled a zero trajectory (Rundblad & Annaz, 2010b). However, in the cross-sectional studies by Chahboun et al. (2016, 2017), an improvement with age was found in high cognitive and verbal functioning individuals with autism.

As Kalandadze et al. (2018) suggests, a novel and useful methodological approach to examine whether figurative language—and, particularly, metaphoric language—is developmentally
delayed or deviant in children with ASD involves including one ASD group and two TD comparison
groups: one matched on CA and another matched on LA. Comparing the three groups will make it
possible to clarify the greater or lesser weight of factors related to age, linguistic competence, or the
ASD itself in the comprehension of metaphors, which is usually associated with the presence of
extralinguistic differences.

**Conventional and novel metaphors**

Not all metaphors require the same level of processing (at linguistic, contextual, and cognitive
levels). According to the Career of Metaphor Model (Bowle & Gentner, 2005), metaphors can vary in
terms of conventionality/novelty. Conventional metaphors (or lexicalized) are familiar, frequent, and
prototypical. Because they are stored in the mental lexicon, their understanding requires a lexical
retrieval process. They have a fixed meaning that has to be retrieved from stored knowledge in
memory. By contrast, novel metaphors are new, original, creative, not familiar, and infrequent. Their
meaning may be discerned without prior familiarity or explanation, requiring a dynamic and creative
process based on the constituent words and the comparison of the concepts they express. Whereas
conventional metaphors are understood through categorization, novel metaphors are understood
through comparison (matching).

The Career of Metaphor Model (Bowle & Gentner, 2005) posits that the comprehension of
conventional metaphors is less demanding, in terms of cognitive and linguistic resources, than the
comprehension of novel metaphors, given that the former, because they are more familiar and
frequent, are processed more quickly and easily than the latter. However, studies carried out with
children comparing the two types of metaphors have obtained inconsistent results. Thus, although
some studies found better performance on the comprehension of conventional metaphors than on
novel ones (Gold, Faust & Goldstein, 2010; Zheng et al., 2015), others obtained similar performance
on both types of metaphors (Kasirer & Mashall, 2016, only in the TD group; Mashall & Kasirer, 2011,
2012; Olofson et al., 2014), and some even found better performance on the comprehension of novel
metaphors than on conventional ones (Chahboun et al., 2017; Kasirer & Mashall, 2016, only in the
ASD group).

In addition, when comparing the groups, children with TD usually obtain better performance
than children with ASD on the comprehension of conventional metaphors (Mashal & Kasirer, 2011,
Minshew, Goldstein & Siegal, 1995). However, in the comprehension of novel metaphors, whereas
some studies also obtained better performance in children with TD (Chahboun et al., 2016; Gold et al., 2010; Olofson et al., 2014; Van Herwegen & Rundblad, 2018), other studies found that the performance on novel metaphors was similar in both groups (Kasirer & Mashall, 2016; Mashall & Kasirer, 2011, 2012; Melogno, D’Ardia, Pinto & Levi, 2012; Zheng et al., 2015).

Objectives of the present study

The main objective of the present study was to evaluate the metaphor comprehension (both overall and depending on the degree of conventionality/novelty) in a sample of children with and without ASD. A novel contribution of our study is that the children with ASD in our sample had level 2 severity, whereas the children included in the samples of almost all the previous studies on figurative language were from level 1, in many cases with high cognitive and verbal functioning. Therefore, our study focuses on a type of population in which figurative language has hardly been studied, and so we address a part of the autism spectrum that differs from what has commonly been studied. Furthermore, in this study, we used a methodological approach that represents an improvement over most of the previous studies because it includes, in addition to the ASD group, two comparison groups: a TD matched on CA with the ASD group but with a higher LA (called TCD), and a TD group matched on LA with the ASD group but with a lower CA (called TLD).

The specific objectives of the present study were the following:

Objective 1. Determine whether there were differences among the three groups of children on the overall performance obtained on the metaphor test.

Objective 2. Determine whether there were differences among the three groups of children on the performance obtained on each of the two types of metaphors (conventional and novel) included in the metaphor test.

Objective 3. Determine whether, within each group, there were differences in the performance obtained on the two types of metaphors (conventional and novel).

Method

Participants

The participants in the present study were 66 elementary school children between 6 and 12 years old, of European Caucasian descent, and with a medium socioeconomic status. They were all native Spanish speakers, and they were divided into three groups: a group of children with an ASD diagnosis (ASD group), a group of children with TD matched with the ASD group on CA (TCD group),
and a group of children with TD matched with the ASD group on LA (TLD group). To assess the LA, the Peabody test (Dunn, Dunn & Arribas, 2006) was used in all cases.

- The ASD group was composed of 18 males and 4 females, with a mean CA of 10.92 years (SD = 1.19) and a mean LA of 8.28 years (SD = 1.88).

- The TCD group was composed of 16 males and 6 females, with a mean CA of 10.89 years (SD = 1.02) and a mean LA of 11.48 years (SD = 1.55).

- The TLD group was composed of 16 males and 6 females, with a mean CA of 8.35 years (SD = 0.32) and a mean LA of 8.64 years (SD = 2.02).

No statistically significant differences in CA were found between the ASD and TCD groups \([F(1,42) = .011, p = .916, \eta^2_p = .001]\), or in the LA between the ASD and TLD groups \([F(1,42) = .375, p = .543, \eta^2_p = .001]\). Moreover, no statistically significant differences were found between the CA and LA in the TCD group \([F(1,21) = .40, p = .056, \eta^2_p = .163]\) or in the TLD group \([F(1,21) = 1.59, p = .221, \eta^2_p = .070]\). Finally, no statistically significant differences were found among the three groups on gender \((\chi^2 = 0.518, p = .472)\) or performance IQ \([F(2,63) = .57, p = .947, \eta^2_p = .002]\), with a performance IQ (evaluated with the Raven test) above 85 in all cases.

Children in the ASD group had a clinical diagnosis of ASD, according to the DSM-5 (APA, 2013), and they met the diagnostic criteria for level 2. The diagnosis was confirmed using more specific instruments, such as the Autism Diagnostic Observation Schedule (ADOS; Lord, Rutter, DiLavore & Risi, 2000), which was applied by specialized psychologists who had official accreditation to use this instrument. In addition, in all cases, the scores obtained on the Social Communication Questionnaire (SCQ; Rutter, Bailey, & Lord, 2003) were above 15 (M = 26.07; SD = 7.08). All the children in the ASD group attended schools with specific classrooms where the TEACCH (Treatment and Education of Autistic and Related Communication Handicapped Children) methodology was used. These classrooms, which are included in regular state schools in XXXX, are set up for children with disorders that affect language and communication. In the case of the children in the two comparison groups, they were children with TD without any clinical diagnosis who attended the same schools as the ASD group, but in the regular modality.

Procedures
The tests were administered in the school, individually, on various days, in a noise- and distraction-free office. In all cases, the tasks were administered in the same order (first: the Peabody test, second: the metaphors test, which was administered on the computer).

Ethics Statement

This study was approved and funded by the University of XXXX and the XXXX (Helsinki Declaration in the Convention of the European Council, 1964). The research had the official and written authorization of the General Direction and School Management (XXXX Education, Training and Employment Department). Moreover, written authorization to carry out the research was obtained from the participating schools and the children's parents.

Instruments

Peabody Picture Vocabulary Test (PPVT-III, Dunn et al., 2006). The Peabody test is an instrument widely used to assess vocabulary knowledge in children. It consists of 192 items: the examiner names a word (noun, verb, adjective, etc.), and the child has to point out an image from four images presented. We used the equivalent linguistic age score provided by the test. The Peabody test is an appropriate test to administer to handicapped individuals with articulation or expressive language problems (Bell, Lassiter, Matthews & Hutchinson, 2001), which is the case of the participants with ASD in our sample. Moreover, in some studies, the validity of the PPVT has been shown by strong correlations between PPVT scores and overall intelligence (Bee & Boyd 2004; Bell et al. 2001), even though it has been used in several investigations on ASD to obtain an estimation of language ability (e.g. Hala, Pexman & Glenwright, 2007; Lam & Yeung, 2012; Pellicano, 2010; Pring, Ryder, Crane & Hermelin, 2010).

Metaphors Test. This instrument, which has a computerized format, includes a total of 20 items that can be classified according to their degree of conventionality/novelty (see annex 1). Thus, half the items were assigned to the category of conventional metaphors (e.g. “This child is a pig”, “This school is a prison”), and the other half were assigned to the category of novel metaphors (e.g. “This girl is a lynx”, “My brother is brilliant”). The following procedure was used to place the items in the categories (conventional / novel): after consulting documental collections of metaphors in Spanish, works of literature, and the media, two experts in Linguistics selected a total of 60 metaphors. A pilot study was carried out in two phases with a total of 225 children with TD between 6 and 12 years old who were not part of the sample of the present study. In the first phase, the children
were asked to indicate whether or not they knew what each of the 60 expressions meant. The 10 expressions that obtained the highest percentages of "yes" responses and the 10 expressions that obtained the highest percentages of "no" responses were selected for the construction of the test. Each of these 20 metaphors was designed as an item. Each item was composed of a statement and three possible response options: one option was correct (the figurative target meaning of the expression), another option approached the literal meaning of the expression, and another option was related to the expression but was not its target meaning. The subject had to point to or indicate which of the three response options fit best or corresponded to the sentence in the item. For example, for the expression “The boy is sunk”, the response options were: (a) The boy is sad; (b) The boy fell; and (c) The boy is little clumsy. To confirm the frequency or familiarity of the items, in the second stage of the pilot study, the 20 items were administered to the 225 children. Based on the frequency of hits obtained on each item, the items were classified in the two categories: conventional (high frequency of hits, above 70%) and novel (low frequency of hits, below 30%). We think this procedure is appropriate because the metaphors that present a high frequency of comprehension are familiar, almost automatized, and stored in the memory, whereas the metaphors that present low frequency of comprehension are new, not familiar, and not stored in the memory. Moreover, because a large number of children with ages similar to those in our sample are used –and not adults- we think the validity of our procedure is strengthened, given that the criterion of frequency is not based on the competence of a group of adults, but rather on that of a group of children. Thus, many metaphors that might be conventional (due to their high frequency or familiarity) for adults might not be for children the first times they are presented because they are novel to them (with low frequency or familiarity). When administering the test to the children in the sample, the test items were presented randomly. For each subject, the number of correct answers (the total number and those obtained on each type of metaphor) was recorded, as well as the time taken to perform the test. Previously, training on the task was provided, administering various items that were different from those on the test in order to make sure the subject understood the task.

Data Analysis

Data analyses were performed using the SPSS version 24 statistical package. For the first objective, two ANOVAs were conducted (one for the number of correct answers, and the other for the response time) to compare the overall performance of the three groups of children on the metaphor
test. For the second and third objectives, a 2 x 3 mixed variance analysis was conducted, including as independent variables: type of metaphor (intra-subject variable) and group (inter-subject variable), and as dependent variable, the number of correct answers. The overall effect of each independent variable and the interactions between them were analyzed, in order to study group differences in the performance obtained on each type of metaphor (objective 2), and within-group differences in the performance on the two metaphor types (objective 3). Effect sizes were calculated using partial $\eta^2$ values, according to Cohen (1988): $\eta^2 = .06$, small effect size; $\eta^2 = .06$ to .14, medium; $\eta^2 = .14$, large.

**Results**

**Objective 1**

Group differences in the overall performance on the metaphor test.

Table 1 presents the results of the ANOVAs conducted to compare the overall performance of the three groups of subjects on the metaphor test, along with pairwise comparisons. The results showed statistically significant differences among the three groups. The TCD group showed the best performance on both the number of correct answers and the response time, whereas the ASD group showed the worst performance.

**Objectives 2 and 3**

Table 2 presents the means and standard deviations for the groups' performance on each type of metaphor (conventional and novel) included in the metaphor test. The mixed ANOVA showed a significant overall effect for the group ($F_{(2,63)} = 39.47, p < .001, \eta^2_p = .556$) and the type of metaphor ($F_{(1,63)} = 128.47, p < .001, \eta^2_p = .671$). Overall, the TCD group showed the best performance, whereas the ASD group showed the worst performance, and more correct answers were obtained on conventional metaphors than on novel metaphors. Moreover, the mixed ANOVA showed a significant overall effect of the interaction between the two independent variables (type of metaphor x group) ($F_{(2,63)} = 6.85, p = .002, \eta^2_p = .179$).

**Regarding objective 2 (analyze group differences in the performance obtained on each type of metaphor), statistically significant differences were obtained among the three groups on both novel ($F_{(2,63)} = 35.30, p < .001, \eta^2_p = .528$) and conventional metaphors ($F_{(2,63)} = 35.32, p < .001, \eta^2_p = .529$). Bonferroni post-hoc pairwise comparisons revealed that, on each of the two metaphor types, the TCD
group obtained more correct answers than the TLD group, and the TLD group obtained more correct answers than the ASD group, with statistically significant differences in all cases ($p$ values ranging from .01 to <.001).

Regarding objective 3 (analyze within-group differences in the performance on the two metaphor types), statistically significant differences were obtained between the two types of metaphors within the three groups (ASD group: $F_{(1,63)}= 42.82, p < .001, \eta^2_p = .405$; TCD group: $F_{(1,63)}= 15.41, p < .001, \eta^2_p = .197$; and TLD group: $F_{(1,63)}= 83.93, p = .571$). In all cases, more correct answers were obtained on conventional metaphors than on novel metaphors (see Figure 1).

**Discussion**

The main objective of the present study was to evaluate the metaphor comprehension (both overall and depending on the degree of conventionality/novelty) in a sample of children with and without ASD. To do so, a metaphor test was constructed that included conventional and novel metaphors, based on the criterion of frequency / familiarity, as in recent studies (e.g., Chahboun et al., 2017).

With regard to the first objective of the present study, the TCD group showed better performance on the test than the TLD group, and the latter obtained better performance than the ASD group. First, the differences found between the two TD groups indicate the importance of the development of both linguistic and extralinguistic skills in metaphor comprehension. Second, the fact that the two TD groups performed better than the ASD group would coincide with previous studies (see the metanalysis by Kalandadze et al., 2018) that found better performance on metaphor comprehension in children with TD than in children with ASD. Third, the differences found between the ASD and TCD groups (in favor of the TCD group) which were matched on CA but not on LA, would reinforce the critical role of core language skills (grammar and semantics) in the comprehension of metaphors (Norbury, 2005; Pouscoulous, 2014). Fourth, the differences found between the ASD and TLD groups (in favor of the TLD group) which were matched on LA but not on CA, would reinforce the idea that other skills and competencies, apart from linguistic skills, would also be involved in the comprehension of metaphors (Vulchanova et al., 2015). Thus, future studies would have to investigate whether difficulties in mentalist skills and social comprehension, executive dysfunction, weak central coherence, and difficulties in integrating the information in a context,
aspects that are commonly found in children with ASD, would be the basis of the limited metaphor comprehension found in this group of children. They would fail to integrate the necessary information, the speaker’s intent, and the rest of the context where all of this must be used (Vulchanova et al., 2015). The context seems to play a key role in metaphor comprehension, which could justify the difficulties in comprehending metaphors in ASD based on recent theoretical approaches such as context blindness theory (Vermeulen, 2009). The possible influence of all these extralinguistic competencies on the comprehension of metaphors in children with level 2 ASD is a topic that should be addressed in future investigations.

The second and third objectives of the present study address the differentiation between conventional and novel metaphors. The comparison of the three groups (objective 2) on the comprehension of each type of metaphor yielded similar results to those obtained on the overall performance on the test. Thus, on both conventional and novel metaphors, the TCD group obtained better performance than the TLD group, which obtained better performance than the ASD group. First, the differences found between the two TD groups, in favor of the older ones, would be explained by the effect of age and the corresponding development of linguistic and extralinguistic skills.

Second, the differences found between the ASD group and the two TD groups would agree with the literature showing better performance in TD than in ASD on both conventional and novel metaphors (Chahboun et al., 2016; Gold et al., 2010; Olofson et al., 2014; Van Herwegen & Rundblad, 2018). However, in the case of novel metaphors, our results would not agree with previous literature that found similar performance in the two groups (Kasirer & Mashall, 2016; Mashall & Kasirer, 2011, 2012; Melogno, D’Ardia, Pinto & Levi, 2012; Zheng et al., 2015).

In the case of conventional metaphors, the results seem conclusive, in that the children with TD perform better than the children with ASD on this type of metaphor. The comprehension of conventional metaphors (familiar, frequent) is carried out through categorization, requiring an information recovery process from long-term memory. This process—which is mainly guided by concepts, or top-down—requires the selection of the relevant or appropriate meaning of the term contained in the expression and the suppression of meanings that are irrelevant or inappropriate. It is a global process in which children with ASD would present greater limitations than children with TD, partly due to weak central coherence (Happé & Frith, 2006; Vulchanova et al., 2015), which could
somewhat justify the worse performance of the ASD group, compared to the TD group, on conventional metaphors.

In the case of novel metaphors, because they are original and infrequent, their comprehension would not require recovering information from long-term memory; instead, it would involve carrying out a comparison in which the semantic features of the base term and the target term are extracted and matched to each other, using their attributes to establish the basis for the comparison. In this case, the necessary processing is mainly guided by data, or bottom-up. It is a local processing (or focused on the details), in which children with ASD can present similar performance to children with TD, according to the theories of weak central coherence and enhanced perceptual functioning (Happé & Frith, 2006; Mottron, Dawson, Soulières, Hubert & Burack, 2006), which could partly justify the lack of differences on novel metaphors found between the groups in some studies. However, we think this lack of differences would be more likely in ASD with high cognitive and verbal functioning, which is not the case in our study (ASD level 2), where higher performance was obtained on novel metaphors in the TD groups than in the ASD group.

The within-group comparison (objective 3) of the two types of metaphors yielded similar results in the three groups: performance on conventional metaphors was better than on novel metaphors. This result would be consistent with some previous studies (Gold et al., 2010; Zheng et al., 2015), and it would support the Career of Metaphor Model (Bowle & Gentner, 2005). In this model, the comprehension of conventional metaphors is considered less demanding, in terms of cognitive and linguistic resources, than the comprehension of novel metaphors, given that the former, due to being more familiar and frequent, do not produce metaphoric tension and are processed more easily than the latter. However, other studies obtained different results (Chahboun et al., 2017; Kasirer & Mashall, 2016; Mashall & Kasirer, 2011, 2012; Olofson et al., 2014), explained by the degree of transparency of the types of metaphors: thus, whereas conventional metaphors are more opaque and lexicalized, and their meaning is difficult to infer without a supportive context, novel metaphors are transparent, and their meaning may be readily discerned without prior familiarity or explanation. Nevertheless, because conventional metaphors have been defined and operationalized as frequent and familiar, learning would play a key role in the better performance obtained on conventional metaphors than on novel ones. Hence, conventional metaphors, as highly familiar expressions stored in the memory, would be better understood than novel metaphors because the access to their
meaning based on the term contained in the expression would be practically automatic. However, this would not occur with novel metaphors because when a novel metaphor becomes frequent or familiar, it turns into a conventional metaphor.

Conclusions

The results obtained from the sample of children in our study lead us to draw a series of conclusions, which, in some cases, form hypotheses to evaluate in future studies. It can be concluded that the understanding of the metaphor is affected by the individuals' language skills, with subjects with ASD (level 2) being more affected than subjects with TD, even when they present the same linguistic age. This suggests understanding metaphors requires other linguistic skills (grammar and semantics) and extralinguistic skills, such as mentalist skills and social comprehension, executive skills, central coherence, integration, and contextual processing skills that are usually limited and compromised in children with ASD, as other studies have pointed out (Baron-Cohen, 1989, 2000; Happé, 1993; Jolliffe & Baron-Cohen, 1999, 2000; Landa & Goldberg, 2005). These skills are even more necessary in the case of metaphors that are new to all the groups studied, although the subjects with ASD have the greatest difficulties in this regard. Moreover, the degree of the metaphor's frequency or familiarity would be a key factor in its comprehension, so that highly frequent and familiar metaphors would be stored in the long-term memory and understood almost automatically, with little mental effort, unlike infrequent metaphors.

Limitations and future research

Our study has some limitations. First, the specific characteristics of the group of children with ASD in our sample limits the reach of the results obtained to only these types of children (ASD level 2, between 6 and 12 years old). Thus, the autism spectrum was not completely represented because there were no children in the sample with levels 1 and 3. Second, regarding sex, the sample used in the present study includes more men than women. Third, because the study is cross-sectional, the variables were not studied over time. In the specific case of the children with ASD, by including only one group, it was not possible to study the evolution of metaphor comprehension with age. Given that contradictory results have been found in this aspect in the scientific literature (e.g. Chahboun et al., 2016, 2017; Rundblad & Annaz, 2010b; Van Herwegen & Rundblad, 2018), this issue should be addressed in future longitudinal studies. Fourth, this study did not have a comparison group with a
different psychological disorder, and so we cannot conclude that the differences found compared to the TD groups were only attributable to the condition of autism. Fifth, in the evaluation of linguistic ability, only one aspect of language was taken into account, specifically, receptive vocabulary. Therefore, it would be necessary to carry out studies that include other aspects of language and their involvement in figurative language comprehension by subjects with ASD. Sixth, future research will have to study -in children with ASD with the characteristics of our sample- the comprehension of other types of figurative language (e.g. irony, metonymy…).

As mentioned above, although it is true that most of the articles published on figurative language in children with ASD have used samples with level 1 severity and high cognitive and verbal functioning, very few studies on the topic have included children with level 2 ASD. In this regard, our study contributes to the research on the topic by filling a gap in the scientific literature and providing data on metaphor comprehension in children with ASD with these specific characteristics.

Acknowledgments

The authors thank the families and the schools for their participation in this research, and the XXXXXXX for providing the required financial resources.

Declaration of Conflicting Interests

The authors declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Funding

This study received human and financial resources from the xxxxx Government (Grant xxxxxxxx).
References


Annex 1. Items included in the metaphor test (translated into English).

What does it mean? Choose the correct response in each case.

(N) This girl is a fox.
   1) This girl has a fox in her house.
   2) This girl is very sly and alert.
   3) This girl likes to play chess.

(C) This boy is a pig.
   1) This boy has pigs in his house.
   2) This boy likes to eat ham.
   3) This boy gets very dirty.

(N) My brother is brilliant.
   1) My brother is very smart.
   2) My brother has a shiny face.
   3) My brother talks a lot.

(N) I’m up to my eyeballs these days.
   1) I’m very busy.
   2) My eyes hurt.
   3) My eyes are blue.

(C) This school is a prison.
   1) I have a hard time in this school.
   2) There are books in this school.
   3) I have a good time in this school.

(C) The news lifted me up.
   1) The news made me happy.
   2) The news made me raise my arms.
   3) The news made me sad.

(C) Pepe has a high position in the company.
   1) Pepe lives on a high floor.
   2) Pepe’s job is very important.
   3) Pepe is very fat.
(N) Juan is very dense today.
   1) Juan isn’t hungry.
   2) Juan has irritated skin.
   3) Juan doesn’t understand things well.

(C) I love Maria’s velvety voice.
   1) Maria is wearing a scarf on her neck.
   2) Maria is hoarse from talking so much.
   3) Maria has a pleasant voice.

(C) My friend has a screw loose.
   1) When they made my friend, they forgot to put a screw in.
   2) My friend does a lot of silly and stupid things.
   3) My friend’s computer is missing a screw.

(N) The classroom was a zoo.
   1) The children were working quietly.
   2) The teacher was a monkey.
   3) The children were running and playing.

(N) The child is down in the dumps.
   1) The child fell down.
   2) The child is sad.
   3) The child is a little clumsy.

(C) Luis doesn’t grasp the idea.
   1) Luis can’t hold it in his hand.
   2) Luis doesn’t understand the explanation.
   3) Luis dropped something.

(N) Quique was feeling low.
   1) Quique was sitting on the floor.
   2) Quique was touching the ground.
   3) Quique was sad.

(N) During the argument, Pedro really attacked Maria.
   1) Pedro was very critical of Maria.
2) Pedro hit Maria.
3) Pedro didn't want to argue with Maria.

(C) Juan wasted his time.
   1) Juan lost a valuable watch.
   2) Juan used his time for something useless.
   3) Juan couldn't find something valuable he had lost.

(N) The conversation with Rachel flowed.
   1) Rachel drools a lot when she talks.
   2) Rachel is a good swimmer.
   3) Rachel expresses herself well when she talks.

(C) Time is golden.
   1) Time goes by very slowly.
   2) Time is important.
   3) The weather is nice today.

(N) Antonio lost the thread of the conversation.
   1) Antonio doesn't know what the conversation was about.
   2) Antonio could not sew a type of clothing.
   3) Antonio spoke in a very soft voice.

(C) Ana had high hopes.
   1) Ana had great hopes and expectations.
   2) Ana shot an arrow high in the air.
   3) Ana doesn't know how to hold a shotgun.

Note: Conventionality (C); Novelty (N)
Table 1.

*Group differences in the overall performance (correct answers and response time) on the metaphor test.*

<table>
<thead>
<tr>
<th></th>
<th>ASD</th>
<th>TCD</th>
<th>TLD</th>
<th>$F_{2,63}$</th>
<th>$\eta^2$</th>
<th>Differences between groups</th>
</tr>
</thead>
<tbody>
<tr>
<td>Correct answers</td>
<td>M</td>
<td>7.68</td>
<td>17.50</td>
<td>12.68</td>
<td>37.57**</td>
<td>.543</td>
</tr>
<tr>
<td></td>
<td>SD</td>
<td>4.30</td>
<td>1.84</td>
<td>4.52</td>
<td></td>
<td>ASD &lt; TLD, TCD; TLD &lt; TCD</td>
</tr>
<tr>
<td>Response time</td>
<td>M</td>
<td>512.13</td>
<td>240.36</td>
<td>357.22</td>
<td>23.93**</td>
<td>.431</td>
</tr>
<tr>
<td></td>
<td>SD</td>
<td>189.11</td>
<td>61.11</td>
<td>107.47</td>
<td></td>
<td>ASD &gt; TLD, TCD; TLD &gt; TCD</td>
</tr>
</tbody>
</table>

**$p<.01$**
Table 2.

*Means and standard deviations obtained by the groups in their performance on the two metaphor types (conventional and novel) included in the metaphor test.*

<table>
<thead>
<tr>
<th></th>
<th>ASD</th>
<th>TCD</th>
<th>TLD</th>
</tr>
</thead>
<tbody>
<tr>
<td>Correct answers on conventional</td>
<td>M</td>
<td>5.22</td>
<td>9.63</td>
</tr>
<tr>
<td>metaphors</td>
<td>SD</td>
<td>2.36</td>
<td>.95</td>
</tr>
<tr>
<td>Correct answers on novel metaphors</td>
<td>M</td>
<td>2.72</td>
<td>8.36</td>
</tr>
<tr>
<td></td>
<td>SD</td>
<td>2.27</td>
<td>1.09</td>
</tr>
</tbody>
</table>
Figure 1. Within-group differences in the performance on the two metaphor types (conventional and novel) included in the metaphor test.