Longitudinal associations between metaphor understanding and peer relationships in middle childhood

Paola Del Sette¹ | Luca Ronchi¹ | Valentina Bambini² | Serena Lecce¹

¹Department of Brain and Behavioural Sciences, University of Pavia, Pavia, Italy
²Center for Neurocognition, Epistemology and Theoretical Syntax (NEtS), School for Advanced Studies IUSS, Pavia, Italy

Correspondence
Paola Del Sette, Department of Brain and Behavioural Sciences, University of Pavia, Piazza Botta 6, Pavia 27100, Italy.
Email: paola.delsette01@universitadipavia.it

Abstract
Metaphor understanding is a pragmatic inferential skill that serves a social function in adulthood. The present longitudinal study was designed to investigate the existence and the direction of the associations between metaphor understanding and peer relationships in children. One-hundred twenty-six typically developing 9-year-old children were tested at baseline and a year later for their ability to understand metaphors (via the physical and mental metaphors task) and their peer relationships (through peer acceptance and rejection). Results showed a longitudinal and bidirectional association between metaphor understanding and peer rejection, but not peer acceptance. Children who were more rejected were less able to understand metaphors 1 year later and children who struggled in understanding metaphors were more rejected 1 year later. Results are discussed in light of existing theoretical models. Overall, these findings show that the social nature of metaphor traces back to peer relationships in children.

Highlights
- The study examined the developmental associations between metaphor understanding and social peer relationships.
Adopting an autoregressive cross-lagged modelling approach we found a bidirectional developmental association between metaphor accuracy and peer rejection. High rejection predicts lower metaphor understanding 1 year later and low metaphor understanding predicts higher rejection 1 year later.

**KEYWORDS**
experimental pragmatics, longitudinal design, metaphor, peer rejection, pragmatics, social relationships

### 1 | INTRODUCTION

Metaphor understanding involves a high-order form of pragmatic inference that allows to fill the gap between the literal and the intended meaning of a sentence conveyed by a social partner (Carston, 2012; Sperber & Wilson, 2002). Different theoretical accounts in the fields of pragmatics, computational and cognitive linguistics have shed light on the processes underpinning metaphor understanding (Carston, 2010; Gibbs, 2011; Glucksberg, 2003; Kintsch, 2001; Sperber & Wilson, 2012; Wilson, 2011). According to the pragmatic account (e.g., Relevance Theory), metaphor understanding requires the listener to adjust the literal meaning of words and to infer the speaker’s intended meaning based on the communicative and social context (Carston, 2010; Sperber & Wilson, 2012). For instance, to understand a metaphor such as ‘Sally is a chameleon’ (from Carston, 2012), the listener is engaged in the fine-tuning of the meaning of the metaphor’s vehicle (i.e., the word ‘chameleon’) by promoting some of its properties (e.g., the characteristics of fitting to the surroundings) and deriving the meaning that the speaker intended to convey in that specific context (for instance, ‘Sally will easily adapt to the new situation’). Differently, in the computational field, researchers have highlighted the role of semantic knowledge in metaphor understanding. For example, according to computational accounts of metaphors such as the predication theory, the meaning of nominal metaphors in the form ‘X is Y’ can be derived from the semantic distance between the topic and the vehicle of the metaphors (Kintsch, 2001; Kintsch & Bowles, 2002). For instance, to derive the meaning of a nominal metaphor such as ‘My lawyer is a shark’, the semantic neighbours of the vehicle that are most closely related to the topic should be selected and used to assign the right meaning to the vehicle (e.g., the semantic neighbours of ‘shark’, such as ‘vicious’, ‘mean’, or ‘aggressive’, are emphasized, whereas the features of ‘shark’ that are not related to the ‘lawyer’, such as ‘has fins’ or ‘swims’, are de-emphasized). This theory is highly compatible with the categorization account, according to which metaphors create novel categories that are used to describe the topics of interest based on outstanding exemplars of those categories (for instance, ‘shark’ is used to build a category of any predatory being), and as such are understood via categorization processes (Glucksberg, 2003; Glucksberg & Keysar, 1990). Finally, in the framework of cognitive linguistics, metaphor understanding is reached through the retrieval of the conceptual mapping across cognitive domains (Gibbs, 2011; Lakoff & Johnson, 1980). Thus, for the metaphor ‘Love is a journey’, entities in the source conceptual domain JOURNEY (e.g., travellers, vehicle, destinations, obstacles, and so on) are mapped onto entities in the target conceptual domain LOVE (e.g., lovers, relationships, common goals, problems, and so on; Gibbs, 2011).

Despite the diverse theoretical accounts on the processes underlying metaphor understanding, there is a general agreement regarding the ubiquitous presence of metaphors in people's daily life. Metaphors are frequently used in everyday conversations, schoolbooks, academic texts, literature and media communications (Cameron, 2008; Glucksberg, 1989; Golden, 2010; Steen, Dorst, & Hermann, 2010) as they serve a social function. On the one hand, metaphors make social communication more persuasive (Landau, Meier, & Keefer, 2010; Sopory & Dillard, 2002) and the interlocutor more attractive (Gao et al., 2017). On the other hand, communicative exchanges involving
metaphors require the activation of shared social knowledge and increase participants’ sensitivity to social cues during conversations (Bowes & Katz, 2015). Thus, metaphors create intimacy by drawing the listener and the speaker closer to one another through the creation of a common ground and a shared understanding (Bowes & Katz, 2015; Cohen, 1978; Horton, 2007). Further support to the social function of metaphors, comes from research on clinical populations showing that the impairment of metaphor comprehension hinders the individual’s social functioning and relationships (Adamczyk et al., 2016; Bambini et al., 2020).

The present study builds on this literature and investigates the existence and direction of the associations between metaphor understanding and children’s peer relationships in middle childhood. Before discussing in more depth each of these main issues, it is important to underline the relevance of such a study in the light of the robust literature showing how peer relationships are crucial for children’s cognitive development and well-being both in the short and in the long term (Rubin, Bukowski, & Parker, 2006). Moreover, the results of the present study will help to shed light on the social origins of important communicative phenomena, such as metaphors, and on the communicative basis of positive peer relationships.

1.1 Peer relationships and metaphor understanding and in middle childhood

The associations between social interactions and metaphor comprehension have been so far typically studied in adults, with the consequence that very little is known on the role of metaphor understanding in children’s social life. This lack of knowledge is especially striking when referred to middle childhood because this period is crucial for the development of both social relationships and metaphor understanding. Regarding social relationships, research has clearly shown that children’s social life changes dramatically as they enter primary school in terms of the amount and complexity of social relationships (Hughes, 2016). While social interactions involving peers are 10% of the total amount of social interactions in 2-year-olds, in middle childhood they become more than 30% (Rubin et al., 2006). Moreover, social exchanges with peers commonly involve multiple partners, take place in several different settings and include complex activities such as conversing, ‘gossiping’ and playing games with or without formal rules (Kuttler, Parker, & La Greca, 2002; Rubin et al., 2006; Zarbatany, Hartmann, & Rankin, 1990). It is also important to note that, during middle childhood, children’s social exchanges are less supervised by adults than before. For this reason, peer relationships become increasingly demanding being more reliant on children’s own cognitive and social skills (Rubin et al., 2006). Notably, important changes also occur in the pragmatic domain during the primary school years, with higher-level pragmatic skills undergoing developmental progression over middle childhood (Bernicot, Laval, & Chaminaud, 2007; Filippova & Astington, 2008; Lecce, Ronchi, Del Sette, Bischetti, & Bambini, 2019; Noveck, Bianco, & Castry, 2001). In particular, metaphor understanding is a sophisticated ability and, as such, its development continues throughout middle childhood with a key turn-point at 9–10 years of age (Lecce et al., 2019; Noveck et al., 2001). While young children struggle in articulating the link between the topic and the vehicle of a metaphor, especially when tested with verbal explanation tasks (Kalandadze, Bambini, & Næss, 2019), their performance significantly improves by the age of 10 years (Winner, Rosenstiel, & Gardner, 1976). Given these changes in social life and higher-level pragmatic skills, middle childhood represents a key developmental stage to investigate the interplay between peer relationships and metaphor comprehension and to trace back the social origin of metaphor as a communicative tool as well as the communicative basis of social relationship. Thus, focusing on metaphor in this developmental period allows to investigate the potential implications of individual differences in metaphor understanding on children’s social relationships at school and, vice versa, the potential role of social relationships as a predictor of advancements in children’s metaphor understanding.

1.2 The existence of an association between pragmatics and peer relationships

Indirect empirical support to the idea that metaphor understanding is associated with children’s peer relationships come from studies examining the association between social relationships and other aspects of
pragmatic abilities, such as children's conversational skills (e.g., the ability to contribute to the communication with relevant comments, to ask questions, and to regulate turn-taking). Typically, these studies adopted direct measures of peer relationships (e.g., sociometric peer nominations) and found an association between children's conversational abilities and their positive social relationships (Black & Hazen, 1990; Black & Logan, 1995; Hazen & Black, 1989; Nærland, 2011; van der Wilt, van der Veen, van Kruistum, & van Oers, 2018, 2019, 2020; van der Wilt, van Kruistum, van der Veen, & van Oers, 2016).

The great majority of research in this area focused on pre-schoolers and the few studies on school-age children have reported controversial results. While some of them found a significant relationship between children's ability to converse effectively and their social position within the peer group (Putallaz & Gottman, 1981; Putallaz & Wasseman, 1989; Rabiner & Gordon, 1992; van der Wilt et al., 2018, 2020), others did not (Bierman & Furman, 1984; Nowicki Jr. & Oxenford, 1989). This inconsistency may lead to think that the relationships between pragmatics and peer relationships are weaker in primary than in preschool years. However, it is important to note that these studies on middle childhood typically assessed communicative skills that are less sophisticated than figurative language understanding and that seem to appear earlier in development (Bohn & Frank, 2019; van der Wilt et al., 2019). For example, preschoolers are already sensitive to discourse pragmatics (Serratrice, 2005) and speech act distinctions (Hübischer & Prieto, 2019). By the age of five, most children can already master some rules of conversation (like the ability to contribute to the communication with relevant comments, to produce fluid utterances, and to regulate turn-taking), and adapt to the interlocutor's needs (Blain-Brière, Bouchard, & Bigras, 2014). It is, therefore, likely that during primary school years variation in these conversational competencies is limited (van der Wilt et al., 2019). Consequently, the focus on conversational competencies in these studies may have reduced the chances to capture individual variation, preventing researchers to detect the social correlates of individual differences in pragmatics during middle childhood. Other studies have explored the link between pragmatics and social relationships using more comprehensive assessments of pragmatics that are also more suitable when working with children in middle childhood and adolescence. Typically, these assessments include items measuring both conversational abilities (e.g., appropriate turn-taking) as well as more sophisticated pragmatics skills (e.g., a good understanding of non-literal language and the use of socially appropriate sentences). These studies found significant associations between children's aggregate scores of communicative ability and children's social life (Leonard, Milich, & Lorch, 2011; Miranda, Berenguer, Rosello, & Baixauli, 2020; Petranovich, Walz, Staat, Chiu, & Wade, 2017; Wolters, Knoors, Cillessen, & Verhoeven, 2013). Those children who have higher pragmatic abilities are more socially competent than those who lag behind in pragmatic development. These findings are promising in that they showed that there might be a significant link between pragmatics and peer relationships in middle childhood. However, in these studies, both constructs (social competence and pragmatics) were gathered by the same source (i.e., parents' report). Thus, the correlation between social competence and pragmatics may have been inflated by a methodological artefact (Podsakoff, MacKenzie, Lee, & Podsakoff, 2003). Whether there is a significant and genuine association between individual differences in pragmatics and in peer relationships during middle childhood remains, therefore, an open issue.

The above considerations represented the starting point of the present study, which adopted a direct measure of a high-level aspect of the pragmatic competence, namely metaphor understanding, and assessed children's peer relationships through a direct measure that taps into both peer acceptance and rejection. The sociometric index of peer acceptance is the number of most-liked nominations received by classmates and reflects how much a child is liked within the class. Differently, the sociometric index of peer rejection refers to the least-liked nomination received by classmates and indicates how much a child is explicitly disliked by peers. Considering the dimensions of peer rejection and peer acceptance distinctively is crucial since these are two aspects that, although correlated, are distinct from one another, with different cognitive correlates (Slaughter, Imuta, Peterson, & Henry, 2015).

1.3 | The direction of the association between pragmatics and peer relationships

The present study adopted a longitudinal design, which allowed us to expand the existing literature and addressed the issue of the direction of the associations between metaphor understanding and social relationships. The few
studies that adopted a longitudinal design mainly considered global pragmatic skills during interactions with peers and adopted indirect measures of pragmatics (i.e., parents’ reports). The results of these studies are interesting as they provided evidence (although indirect) that children’s sociometric status is predicted by pragmatic skills. For example, children who score higher on pragmatic measures are less likely to show social difficulties and more likely to be accepted by peers (Burleson, Delia, & Applegate, 1992; Law, Rush, Clegg, Peters, & Roulstone, 2015). Similar results come also from a quasi-experimental design study by Place and Becker (1991) on third and fourth graders. The authors showed that children liked a female peer more when she displayed appropriate pragmatic skills during a conversation than when she did not. In a similar vein, training studies showed that promoting conversational pragmatic skills in fifth- and sixth-grade children (e.g., responding appropriately to the conversational partner) had a positive effect on their social interactions as assessed through observations (Bierman & Furman, 1984) and on children's peer acceptance as measured via a roster-and-rating sociometric scale (Ladd, 1981).

Although the cited empirical literature suggests that pragmatic skills affect children’s social life, it is important to note that none of the above studies has tested the role of social relationships in the development of pragmatics in middle childhood. Several existing theoretical models are in line with the hypothesis that social relationships prompt pragmatics, emphasising the importance of peer interactions for children’s development. On the one hand, socio-constructivist theories claim that participating in social exchanges with others promotes cognitive development, since social interaction provides the opportunity to practice, learn and co-construct cognitive skills and knowledge (Tudge & Winterhoff, 1993). For example, focusing on the development of high-order pragmatic skills, playing with others allows children to extend the use of conventional terms beyond the literal meaning of words (Clark, 2019).

On the other hand, several models emphasise the role of negative experiences, such as active peer exclusion (Williams & Nida, 2011), in limiting cognitive development through different mechanisms (Baumeister, DeWall, Ciarocco, & Twenge, 2005; Buhs & Ladd, 2001; Muraven & Baumeister, 2000; Muraven, Tice, & Baumeister, 1998; Williams & Nida, 2011). According to these models, negative social experiences limit cognitive functioning by affecting individuals’ social engagement, emotional reactions and cognitive resources. Further support to the role of peer relationships in children’s cognitive development comes from longitudinal studies focusing on peer acceptance and rejection. Research in this field has shown that being highly accepted and/or less rejected by peers influences children’s development of cognitive skills, including social understanding (Banerjee, Watling, & Caputi, 2011), executive functions (Lecce, Bianco, & Ronchi, 2020), as well as cognitive engagement and academic performance (Buhs & Ladd, 2001; Chen, Hughes, Liew, & Kwok, 2010; Fite, Hendrickson, Rubens, Gabrielli, & Evans, 2013; Oberle & Schonert-Reichl, 2013; Wentzel & Caldwell, 1997). Starting from these findings, we hypothesized that the effect of peer relationships on cognitive development could extend to the development of sophisticated pragmatic inferential skills, like metaphor understanding.

2 | THE PRESENT STUDY

The main aim of the present study was to investigate, during middle childhood, the existence and the direction of the developmental associations between children’s peer relationships and metaphor understanding.

Regarding the first aim, available data point to the existence of significant associations between social relationships and pragmatics understanding (van der Wilt et al., 2019). Support, although indirect, to the expectation there is an association also between social relationships and metaphor understanding in school-age children comes from two main sources of evidence. First, research has shown that metaphors serve a social function in adulthood (Bowes & Katz, 2015; Cohen, 1978; Horton, 2007). Second, social relationships and metaphor understanding were found to display a parallel developmental timing with a burst during middle childhood (Rubin et al., 2006; Winner et al., 1976).

Regarding the directionality, we expect to find bidirectional associations between metaphor understanding and peer relationships over time. On the one hand, existing empirical evidence showed that pragmatic skills (considering both conversational pragmatic skills and the ones measured via comprehensive measures of pragmatics) influence
the development of positive social relationships (Bierman & Furman, 1984; Burleson et al., 1992; Ladd, 1981; Law et al., 2015; Place & Becker, 1991). Thus, we expected that this effect extends to high-level pragmatic inferential skills, such as metaphor understanding. In detail, we hypothesized that children with a better understanding of metaphors become less rejected and more accepted by peers. On the other hand, theoretical models and empirical evidence suggested that social interactions promote the development of cognitive skills (Clark, 2019; Leccce et al., 2020; Rubin et al., 2006; Vygotsky, Cole, John-Steiner, Scribner, & Souberman, 1978). Thus, we expected to find that children who are more accepted by peers are more likely to develop a deeper understanding of metaphors. Moreover, given the theoretical accounts and the empirical evidence claiming that facing stressful experiences, including peer rejection, hinders cognitive development (Banerjee et al., 2011; Buhs & Ladd, 2001; Fite et al., 2013; Leccce et al., 2020; Muraven et al., 1998; Muraven & Baumeister, 2000), we also expected that being rejected by peers would threaten the development of metaphor understanding.

To test these expectations (i.e., bidirectional associations between metaphor understanding and peer relationships), we designed a two-wave cross-lagged longitudinal study in which we measured individual differences in children’s metaphor understanding and peer relationships at baseline and 1 year later. Notably, we also measured and accounted for children’s vocabulary skills in our analyses. This was done because empirical literature suggests that receptive language could have a role in the association between social relationships and pragmatics (van der Wilt et al., 2019, 2020). Thus, considering children’s vocabulary allowed us to examine the uniqueness of the associations between metaphor understanding and peer relationships. Measuring vocabulary allowed us to examine also whether individual differences in language are related to metaphor understanding and to contribute to the theoretical debate on the mechanisms underlying metaphor understanding. If we found that vocabulary uniquely predicts later metaphor understanding, this would fit with the computational models of metaphors and the view that semantic distance accounts for metaphorical meanings (Kintsch, 2001; Kintsch & Bowles, 2002). If we found that social relationships contribute to predict later metaphor understanding, this would be more in line with the pragmatic account of metaphors (Carston, 2010; Sperber & Wilson, 2012).

Finally, in addressing the issues of existence and direction of the associations between peer relationships and metaphor understanding we also considered the role of gender, given that van der Wilt et al. (2016) found associations between peer relationships and pragmatics only in boys and not in girls.

3 | MATERIALS AND METHOD

3.1 | Participants

One hundred sixty-one Italian children attending Year 4 of the primary school participated in the present study. Participants were recruited within a bigger longitudinal study designed to investigate social relationships and cognitive development during middle childhood. In the present study, we focused on two time points, 1 year apart, to explore the longitudinal associations between social relationships and pragmatics skills. Among these 161 children, 33 were not eligible due to specific learning disorders, developmental delays or not being native speakers of Italian. All children who took part in at least one time point were included in our analyses. The final sample consisted of 126 typically developing school-age children (67 F), aged around 10 years at T1 (M = 9.10 years, SD = 0.3, age range: 9.2–10.6). Of these 126 children, 112 were tested in all variables at both time points. The percentage of missing values ranged from 1.56% to 8.59% (see Table 1 for detailed sample sizes). The highest rate of missing data observed at T1 was due to some new children recruited at T2. Other missing data (e.g., metaphor understanding at T1) were due to some children being absent from school the day of the data collection. In the modelling analyses, we handled missing data by the full information maximum likelihood estimation (FIML; Enders & Bandalos, 2001).
3.2 | Procedure

We collected parental written consent for all children at the beginning of the study. Children’s metaphor comprehension, peer acceptance and peer rejection were evaluated in collective sessions at baseline (T1, May 2018) and 1 year later (T2, May 2019). At T1 children were also assessed for vocabulary abilities and asked to complete a family affluence questionnaire. This study was approved by the local University Ethical Committee.

3.3 | Measures

Vocabulary. Vocabulary was measured at T1 using the Italian version of the vocabulary subtest of the primary mental abilities (PMA), intermediate form (age 11–17; Rubini & Rossi, 1982; Thurstone & Thurstone, 1962), in which children were asked to find, in eight minutes maximum, the synonyms of 50 target words choosing among five alternatives (range 0–50). Our choice of collecting the intermediate form of the PMA vocabulary task (instead of the basic one indicated for children ranging from 8 to 11) was guided by the findings from a pilot study in which we administered both the basic and intermediate forms to 55 9-year-old children (M = 9 years;10 months, SD = 0.3). Results from this pilot study showed that, while the two forms of the PMA vocabulary task were correlated, \( r = .33, p = .008 \), 23.6% of children scored at ceiling (i.e., 30 out of 30) on the basic form, Mean = 27.82, SD = 2.47, Mode = 29, Q1 = 27, Q2 = 29, Q3 = 29, while scores were more distributed in the intermediate form, Mean = 18.51, SD = 5.60, Mode = 18, Q1 = 14, Q2 = 18, Q3 = 23, with no children scoring lower than 9 out of 30. As a consequence, the distribution of scores for the basic form significantly deviated from normality, \( D(55) = 0.22, p < .001 \), \( W(55) = 77, p < .001 \), Skewness = −2.40. On the contrary, children’s scores on the intermediate form were normally distributed, \( D(55) = 0.07, p = .200 \), \( W(55) = 97, p = .271 \), Skewness = 0.17.

Socioeconomic status. Socioeconomic status was assessed at T1 through the family affluence scale (FAS; Currie et al., 2008). In this questionnaire, children were asked to answer four questions evaluating: (a) family car ownership (range = 0–2), (b) having an unshared bedroom (range = 0–1), (c) the number of computers at home (range = 0–3), and (d) the number of family vacations in a 12-month period (range = 0–3). An overall index of family socioeconomic background was obtained by summing the score for each question and could range from 0 to 9 points. Previous studies showed that FAS is a sound measure for assessing socioeconomic status. Specifically, the high validity of the FAS was proved by the robust associations found between (a) 11-year-old children’s and their parents’ responses (Andersen et al., 2008), and (b) its score and the gross domestic product (Boyce, Torsheim, Currie, & Zambon, 2006).

Peer relationships. Children indexes of peer acceptance and peer rejection were obtained at T1 and T2 through the sociometric peer nomination procedure originally described in Coie, Dodge, and Coppotelli (1982). According to this procedure, each child was asked to nominate up to three classmates he/she liked the most and up to three classmates he/she liked the least. Children were allowed to nominate only peers belonging to their classroom. Cross-gender nominations were permitted. An index of peer acceptance was obtained for each child by summing the number of positive nominations (most-like nomination, ML) that he/she received from peers. Similarly, an index of peer rejection was obtained by summing the total number of negative nominations (least-like nomination, LL). Finally, in order to control for differences in class size, the ML and LL scores were standardized within classrooms. Notably, children who were not eligible for the inclusion in the analyses (see Section 3.1) participated in the nomination procedure (both as nominator and nominee) in order to improve the reliability and the validity of the sociometric measure (Cillessen & Marks, 2017).

Peer nomination is a widely used sociometric technique to index peer relationships in middle childhood that was shown to have high test–retest reliability (Jiang & Cillessen, 2005). Moreover, indices of peer acceptance and peer rejection show construct validity in that variability in the number of positive (ML) and negative (LL) nominations by peers is highly associated with social behaviour (Rubin et al., 2006) and socio-cognitive understanding (Banerjee et al., 2011). Concerning internal consistency, previous research studies consistently found high-reliability scores for
**TABLE 1**  Bivariate and partial correlations (below the diagonal) among study variables and descriptive statistics

| Tasks                              | T1          | T2          | Descriptive statistics |
|------------------------------------|-------------|-------------|------------------------|
|                                    | ML | LL | Acc | ML | LL | Acc | Valid | Mean (SD) | Observed range |
| Family affluence scale (FAS)       | 0.11 | 0.25** | 0.07 | 0.14 | -0.18* | -16 | 120 | 2.72 (0.47) | 1–3 |
| Vocabulary (Voc)                   | 0.03 | 0.02 | 0.32** | 0.12 | -0.10 | 0.24* | 121 | 18.69 (5.51) | 6–34 |
| Most-Like nomination (ML)          | – | -0.42*** | -0.03 | 0.47*** | -0.36*** | -0.036 | 121 | 0.15 (1.00) | -1.62–3.16 |
| Least-like nomination (LL)         | -0.40*** | – | -0.12 | -0.25** | 0.66*** | -0.18* | 121 | -0.04 (0.95) | -1.31–3.89 |
| Metaphor accuracy (Acc)            | -0.07 | -0.12 | – | -0.06 | -0.25** | 0.45*** | 115 | 15.14 (3.81) | 0–21 |
| Most-like nomination (ML)          | 0.46*** | -0.26** | -0.11 | – | -0.39*** | 0.05 | 124 | 0.12 (1.02) | -1.59–3.60 |
| Least-like nomination (LL)         | -0.33*** | 0.68*** | -0.22* | -0.37*** | – | -0.27** | 124 | -0.10 (0.93) | -1.02–3.35 |
| Metaphor accuracy (Acc)            | -0.04 | -0.19* | 0.40*** | 0.03 | -0.27** | – | 123 | 16.72 (3.41) | 0–22 |

Note: Partial correlations controlled for vocabulary and family affluence. Metaphor accuracy is reported after excluding the ceiling item; therefore, the possible range is 0–22.

*+p < .06.*
*p < .05. **p < .01. ***p < .001.*
peer rejection (i.e., around 0.70—Babcock, Marks, Crick, & Cillessen, 2014; Marks, Babcock, Cillessen, & Crick, 2013; van den Berg & Cillessen, 2012) and low to moderate reliability scores for peer acceptance, which commonly range from around 0.10 to 0.40 (Babcock, 2001; Babcock et al., 2014). We evaluate internal consistency scores for peer acceptance and peer rejection in our sample, following the procedure described in Cillessen and Marks (2017) and using the Kuder–Richardson Formula 20 (Kuder & Richardson, 1937). In line with previous literature, our mean reliability scores across classrooms were 0.28 \( (SD = 0.31) \) at T1 and 0.41 \( (SD = 0.39) \) at T2 for peer acceptance, and 0.71 \( (SD = 0.13) \) at T1 and 0.79 \( (SD = 0.06) \) at T2 for peer rejection. Finally, test–retest reliability scores across 1 year were high for both peer acceptance and peer rejection \( (r > .70, ps < .001) \), as shown by correlations (Table 1.) and path analyses (see Section 4 and Figure 1).

**Metaphor understanding.** Metaphor understanding was assessed at T1 and T2 via the 12-item version of the physical and mental metaphors (PMM) task (Del Sette, Bambini, Bischetti, & Lecce, 2020), which is an extended version of the task originally created by Lecce et al. (2019). This version of the PMM task encompasses a set of six physical and six mental non-lexicalized metaphors (see Del Sette et al., 2020 or click the following link, https://osf.io/5aehk/?view_only=bea12875b984f4f799a66886e1ab2a9d3 for a full description of the PMM items). The PMM is a verbal explanation task in which children are asked to explain the meaning of nominal metaphors in the form ‘X is Y’ (e.g., ‘Le ballerine sono farfalle’ ['Dancers are butterflies'], ‘Il papa è un vulcano’ ['Daddy is a volcano']). Children’s answers were coded according to the level of accuracy, indicating their ability to articulate the link between the metaphor’s topic and the vehicle, on a three-point scale distinguishing incorrect, incomplete/non-salient, and complete/salient answers. More specifically, accuracy scores followed the scoring guidelines in Lecce et al. (2019) and Del Sette et al. (2020) and were based on the theoretical assumption that the meaning of metaphors can be fully grasped based on their salience and regardless of contextual information (Giora, 2003). In this view, the most accurate meaning (i.e., the most salient one) of nominal metaphors in a minimal context, such as ‘Soldiers are lions’, can be inferred based on conventionality, frequency/familiarity, and prototypicality (Giora, 2003). The scoring guidelines of the accuracy coding (Del Sette et al., 2020) were validated through an analysis of salience for each item in the PMM. In this analysis, Del Sette et al. (2020) asked 52 university students to verbalize the meaning of the 12 metaphors. For each metaphor, the authors listed the obtained features, calculated their frequency of occurrence and selected the most salient features (i.e., the features that accounted for at least 50% of all features). Overall, the most salient features extracted via the analysis of salience were consistent with the answers receiving a full score (i.e., a score of 2) according to the PMM scoring guidelines. On this basis, the scoring worked as follows: 0 was assigned to incorrect (e.g., for the metaphor ‘I soldati sono leoni’ ['Soldiers are lions'], ‘Saltano’ ['They jump']), literal (‘Cacciano bene’ ['They hunt well']), and ‘don’t know’ answers. A score of 1 was assigned to incomplete answers (‘Sono bravi a combattere’ ['They are good in fighting']) or answers referring to non-salient features of the metaphor’s topic (‘Corrono veloci’ ['They run fast']), whereas the score of 2 was given to answers that were complete and referred to salient features of the metaphor’s topic (‘Sono coraggiosi’ ['They are courageous ‘;‘Sono forti’ ['They are strong’]).

Inter-rater agreement was calculated through Cohen’s kappa on the 25% of the answers, showing an almost perfect agreement \( (k = 0.91 \text{ at T1 and } k = 0.88 \text{ at T2}) \). Coders were blind to the hypotheses of the study. In order to investigate the psychometric properties of the PMM task, we run a series of confirmatory factor analyses (CFA) assessing internal consistency and test–retest reliability. Results (see Appendix) revealed: (a) that all items, except one (that was removed) loaded on a single factor, (b) good internal consistency and strong test–retest reliability. Based on these findings, we computed a total accuracy index by summing the scores on 11 PMM items (range 0–22).

### 3.4 Statistical analyses plan

Before testing our main hypothesis, we conducted the following preliminary analyses: (a) we performed descriptive statistics for all the study measures, (b) we conducted a mixed ANOVA to examine developmental changes in
metaphor accuracy between T1 and T2, as well as possible gender differences within and across time and (c) we calculated the correlation matrix for all the study variables.

To examine the main hypotheses of the study (i.e., the existence and the direction of reciprocal developmental associations between metaphor understanding and peer relationships), we adopted an autoregressive cross-lagged modelling approach (Selig & Little, 2012) and controlled for longitudinal stabilities, vocabulary, and family affluence. Finally, we adopted a multiple-group modelling approach with gender as the grouping variable to test for possible differences in the hypothesized longitudinal associations between peer relationships and metaphor accuracy across boys and girls. Specifically, for all the models, we used a maximum likelihood estimation with robust standard errors (MLR; Brown, 2015; Kline, 2011). The fit of each model was evaluated following the criteria recommended by Brown (2015): a non-significant Chi-square ($\chi^2$) test, a root mean square error of approximation (RMSEA) $\leq 0.08$, a comparative fit index (CFI), and a Tucker–Lewis index (TLI) $> 0.90$. For the nested models, Satorra–Bentler Chi-square difference test (Satorra, 2000; Satorra & Bentler, 2001) suitable for MLR estimator was used.

4 | RESULTS

4.1 | Preliminary descriptive and correlation analyses

Descriptive statistics among all study variables are reported in Table 1.

A mixed ANOVA was conducted on the metaphor accuracy total scores with time as the within-subject variable and gender as the between-subject variable. Results showed a main effect of time, $F(1,110) = 19.79$, $p < .001$, $\eta^2 = 0.15$, but not of gender, $F(1,110) = 0.0003$, $p = .99$, $\eta^2 < 0.001$, and no time by gender interaction effect, $F(1,110) = .005$, $p = .945$, $\eta^2 < .001$. In detail, we found that metaphor accuracy scores increased from T1 to T2.
These results did not change when we controlled for vocabulary, $F(1,109) = 5.76, p = .018, \eta^2 < .05$, $F(1,109) = 0.08, p = .78, \eta^2 = 0.001, F(1,109) = 0.0002, p = .99, \eta^2 < 0.001$.

Table 1 also reports bivariate correlations among all the study measures. Each of the key variables showed moderate to high rank-order stability across time ($r_s > .45, ps < .001$). Regarding within-time correlations, there was a significant negative correlation between metaphor accuracy and LL at T2, but not at T1. ML was not related to metaphor accuracy at any time point.

4.2 | Longitudinal associations between peer relationships and metaphor understanding

To investigate reciprocal associations between peer relationships (i.e., peer acceptance and peer rejection) and metaphor understanding, we specified an autoregressive cross-lagged path-analysis model. More precisely, our model included autoregressive effects connecting the same variables over the two time points, as well as longitudinal paths liking peer relationships (both peer rejection and peer acceptance) and metaphor accuracy in both directions. In the model, we also included socioeconomic status and vocabulary as control variables, predicting all the variables at both T1 and T2. Finally, we permitted error terms to covary within time to allow for potential residual associations among variables due to unmeasured third variables. The resulting model exhibited good fit to the data, $\chi^2(2) = 2.969, p = .23, \text{RMSEA} = 0.06, \text{CFI} = 0.99, \text{TLI} = 0.92$. Results showed that early peer rejection exerted a significant negative effect on later metaphor accuracy, $B = -0.18, p = 0.039, \beta = -0.17$, and that early metaphor accuracy exerted a significant negative effect on later peer rejection, $B = -0.12, p = .046, \beta = -.14$. Peer acceptance did not show any significant association with metaphor accuracy.

When we tested a constrained model in which cross-lagged paths between peer rejection and metaphor accuracy were forced to be equal in both directions, results continued to exhibit a good fit to the data, $\chi^2(3) = 3.42, p = .33, \text{RMSEA} = 0.03, \text{CFI} = 1.00, \text{TLI} = 0.98$. The constrained model showed no significant decrease in the model fit compared to the previous unconstrained model, $\Delta \chi^2(1) = 0.33, p = .565$, indicating that the strength of the regression path from early metaphor accuracy to later peer rejection was equal to the strength of the regression path from early peer rejection to later metaphor accuracy. This final constrained model is presented in Figure 1 with standardized parameter estimates.

Finally, we adopted a multi-group procedure to evaluate possible differences in the cross-lagged associations between metaphor accuracy and peer relationships across gender. Results showed no significant decrease in the model fit when all cross-lagged paths were constrained to be equal across gender, $\Delta \chi^2(4) = 6.20, p = .185$, indicating that the strength of the associations between metaphor accuracy and peer relationships did not differ across gender.

5 | DISCUSSION

The present study was designed to test the existence and the direction of the associations between individual differences in metaphor understanding and in social relationships during middle childhood.

Before discussing the main results, we will comment on preliminary findings concerning developmental changes and individual differences in metaphor understanding. As expected, we found that children’s ability to understand metaphors improved over the 1-year period considered. This finding is in line with the developmental progression observed in previous studies in middle childhood (Del Sette et al., 2020; Pinto et al., 2011; Winner et al., 1976) and indicates that the PMM task is a measure sensitive to developmental changes in middle childhood. Our data also showed that the PMM task was sensitive to capture substantial (between subjects) variation in metaphor understanding in middle childhood. While some children excelled, others struggled in understanding metaphors within a narrow age range (baseline range age 9.2–10.6), with no participants performing at ceiling at any time point.
These findings, together with the one from the CFA (see Appendix), stress that individual differences gathered using this task were genuine and did not reflect transient fluctuations in performance (Bornstein, Putnick, & Esposito, 2017). Overall, the PMM task was found to be a psychometrically sound measure of metaphor understanding in children, with satisfactory statistical properties.

Turning to the main goal, the present study is the first to explore the link between a high-level inferential pragmatic skill, namely metaphor understanding, and social relationships in the critical period of middle childhood. In detail, our aim was to investigate the developmental associations between metaphor understanding and social relationships considering both peer rejection and peer acceptance. Based on empirical evidence and theoretical models presented in the introduction, we hypothesized that children with a better understanding of metaphors become less rejected and more accepted and that children who are more accepted and less rejected are more likely to develop a better understanding of metaphors. Our analyses confirmed the hypothesized bidirectional relationship between peer rejection and metaphor understanding, but not the one between peer acceptance and metaphor understanding. Children with a better understanding of metaphors become less rejected over time and children who are less rejected are more likely to develop a better understanding of metaphors. Notably, these developmental associations were equal for females and males and were independent of longitudinal stability in the key variables (i.e., children’s metaphorical skills and social relationships) and of individual differences in vocabulary and family affluence.

These findings indicate that the negative association between metaphor understanding and peer rejection is not necessarily counterbalanced by the positive one between metaphor understanding and peer acceptance. This fits with existing literature showing that peer acceptance and peer rejection reflect different features of children’s social life (Gifford-Smith & Brownell, 2003; Lecce et al., 2020; Rubin et al., 2006) and point to the detrimental role of the negative experience of being excluded by peers (i.e., peer rejection), rather than to the learning occasions provided by acceptance.

The finding that peer rejection carried the most significant implications for metaphor comprehension suggests that the absence of active rejection by classmates might be sufficient to allow children to develop high-level pragmatic skills across middle childhood. From a theoretical perspective, this claim fits with the view that bad social experiences are stronger than good ones in affecting a range of psychological phenomena, including children’s learning, self-esteem, and cognitive engagement (Baumeister, Bratslavsky, Finkenauer, & Vohs, 2001). Empirical support to this view comes from data showing that, in middle childhood, the absence of active rejection is a key antecedent of good social understanding (Banerjee et al., 2011), cognitive inhibition (Lecce et al., 2020), and attention (Ji et al., 2019), as well as positive self-esteem and school engagement (Danneel et al., 2019). Crucially, all these studies found that peer acceptance did not confer any particular benefit in terms of gains in these cognitive and emotional outcomes (Banerjee et al., 2011). More broadly, our finding that peer rejection negatively predicts later metaphor understanding fits within a large body of literature focusing on the impact of peer rejection on children’s socio-emotional and cognitive functioning (Baumeister et al., 2005; Buhs & Ladd, 2001; Muraven et al., 1998; Muraven & Baumeister, 2000; Williams & Nida, 2011). According to this literature, peer rejection is a major life stressor that threatens the fundamental social need for belonging (Baumeister & Leary, 1995), triggers painful feelings and negative behaviours (Ren et al., 2018), and depletes cognitive resources (Muraven & Baumeister, 2000). As a consequence, peer rejection leads to a variety of negative outcomes (Rubin et al., 2006), including poor academic achievement (Coie, Lochman, Terry, & Hyman, 1992), externalizing and internalizing problems (Burks, Dodge, & Price, 1995; Gazelle & Rudolph, 2004; Ladd & Burgess, 2001; Laird, Jordan, Dodge, Pettit, & Bates, 2001), low classroom participation (Buhs & Ladd, 2001) and poor cognitive development (Banerjee et al., 2011; Fite et al., 2013; Lecce et al., 2020). Our study extends this literature by showing the negative effects of peer rejection on the development of a specific cognitive function, namely metaphor understanding.

Different theoretical models contribute to account for the variety of mechanisms underlying the effect of peer rejection on cognitive outcomes. For example, according to the regulatory-depletion model, individuals’ inner cognitive resources are limited and are depleted by the regulation of negative emotions (frustration and negative affect) derived from stressful life events (Muraven et al., 1998; Muraven & Baumeister, 2000). According to the
need-to-belong model, individuals have a fundamental need to belong and when this need is undermined, they lose motivation to self-regulate and may become unwilling to spend communicative efforts to understand others (Baumeister et al., 2005; Baumeister & Leary, 1995). In a similar vein, the temporal need-threat model claims that when peer rejection is prolonged, it depletes children's coping resources leading to experiences of alienation, depression, helplessness, and feelings of unworthiness (Williams, 2009; Williams & Nida, 2011). Future studies are needed to examine the above-mentioned models in relation to metaphor comprehension and to discriminate the specific processes through which peer rejection impacts metaphor understanding over time.

Interestingly, we found that peer rejection, but not vocabulary skills, uniquely predicted metaphor understanding over time. This finding contributes to the theoretical debate on the processes underlying metaphor understanding (see Section 1) by indicating that social exchanges with peers play a bigger role compared to children's semantic knowledge in predicting metaphor understanding. This pattern of finding fits more with a pragmatic (Carston, 2010; Sperber & Wilson, 2012) rather than a computational account of metaphor (Kintsch, 2001; Kintsch & Bowlles, 2002) since it seems to suggest that the mechanisms through which metaphors are understood cannot be described in terms of semantic processes alone.

Regarding the other direction that is from metaphor comprehension to social relationships, the finding that higher metaphorical skills predicted lower peer rejection over time complements empirical evidence on pragmatic conversational skills in preschool-age children (see, for a review, van der Wilt et al., 2019). Our results nicely fit with the ones reported by van der Wilt and colleagues showing that children with poor pragmatic skills are less attractive playmates and, consequently, become highly rejected (van der Wilt et al., 2018, 2020). Not being able to understand metaphors might represent a risk factor for social relationships in that it makes children less appealing conversational partners. Individuals with poor metaphor understanding skills are limited in their abilities to understand the implicit aspects of discourse, and conversation with them might more easily suffer from breakdowns, besides being more boring: this might ultimately prevent peers from engaging in communicative exchange with those who are poor metaphor comprehenders.

Regarding peer acceptance, the lack of a significant predictive effect on later metaphor understanding was unexpected in the light of socio-constructivist theoretical models (Piaget, 1932; Rubin et al., 2006; Tudge & Winterhoff, 1993; Vygotsky et al., 1978) and certainly requires further investigations. We do not think that positive social interactions, as indexed by peer acceptance, have no effect at all on pragmatic skills. For example, it is possible that social exchanges, and consequently peer acceptance, are relevant for the emergence of other pragmatic skills, such as metaphor production (Clark, 2019). Indeed, the more a child is popular the more the child might feel motivated in using creative expressions such as metaphors, which carry communicative benefits but are also more cognitively challenging for the interlocutor (for the costs and benefits of metaphors, see Noveck et al., 2001). When children are less accepted, they might refrain from more ‘risky’ communicative choices. Our results also suggest that understanding metaphors, although important for preventing children from being actively disliked by peers, may not be sufficient to promote children's likeability. This result seems in contrast with the literature arguing for the social function of metaphors in adulthood (Bowes & Katz, 2015). We think that this is not the case for two main reasons. First, the literature on the social function of metaphor in adulthood mainly adopted a dyadic, rather than a group, approach to social relationships and focused on adults interacting with (just) one social partner. The distinction between dyadic and group relationships is relevant, given that dyadic relationships involve mutual feelings while group relationships as evaluated using the sociometric procedure refer to unilateral perceptions about an individual across the entire peer group (see for a review Flannery & Smith, 2017). Interestingly, the skills that make a child accepted are not necessarily the same skills that make that child an intimate and close partner in dyadic relationships (Flannery & Smith, 2017; Larson, Whitton, Hauser, & Allen, 2007). Second, while in the adult literature metaphors are presented before the evaluation of the closeness to measure the expected prompting effect on social relationships, in the present study, we examined the associations between peer acceptance with the general ability to understand metaphors, measured using an offline task. For these reasons, we do not exclude that metaphor
understanding, when presented before the interaction with a peer, would have a positive effect on that dyadic peer interaction.

6 | CAVEATS AND CONCLUSIONS

The findings of the present study consistently point to a bidirectional relationship between metaphor comprehension and peer rejection (but not acceptance). The study pioneers and leaves some open questions. For example, future research should consider other measures of children's social relationships (e.g., friendship and dyadic likeability) to assess if dyadic relationships, differently from acceptance within the peer group, are associated with metaphor understanding. Such a result is expected since the quasi-experimental design study by Place and Becker (1991) showed that during a dyadic conversation the social partner is more liked when displaying appropriate pragmatic skills. In addition, it would be interesting to complement peers' evaluation of social relationships with teachers' evaluation. In this respect, a recent paper reported a link between pragmatics and children's likeability only when rated by teachers (Cheung & Elliott, 2017), suggesting that peers' and teachers' ratings convey different meanings of likeability that are differently related to pragmatic skills.

Another limitation of the present study regards the control variables. Although we partially addressed this issue by accounting for children's vocabulary skills and SES, a broader range of potential confounding third variables should be included in future studies. For example, it would be interesting to include measures of executive functioning and Theory of Mind, two variables that are known to be associated with both social relationships (Banerjee et al., 2011; Rubin et al., 2006) and pragmatics (Lecce et al., 2019; Matthews, Biney, & Abbot-Smith, 2018) during middle childhood. Future studies should also consider other pragmatic phenomena, such as irony and humour, to examine whether the association between metaphor understanding and peer relationships in middle childhood is specific for metaphors or generalized to other high-level pragmatic skills. In addition, it could be interesting to examine whether other social variables (e.g., prosocial and antisocial behaviour, personality traits, empathy), that are related to children's sociometric status and pragmatics, have a role in the association between pragmatic functioning and social relationships (Andre, Mancini, Mazzoni, Russo, & Baldaro, 2015; Christensen, 2007; de Bruyn & Cillessen, 2006; van den Brink et al., 2010; van der Ploeg, Kretschmer, Salmivalli, & Veenstra, 2017). Finally, future studies should also assess the online production—in addition to the understanding—of metaphors during a direct observational measure of children's social relationships to investigate whether peer acceptance promotes the development of metaphor production, as suggested by Clark (2019).

Granted these limitations, the present study has also important strengths. The first is that it focused, for the first time in literature, on a high-level inferential pragmatic skill, namely metaphor understanding, that is known to have a social role in adulthood (Bowes & Katz, 2015), and whose social features trace back to middle childhood. Second, the present study uses a newly developed task, the PMM, that showed good psychometric properties and can, therefore, be used in future studies. The third strength refers to the use of a comprehensive measure of peer relationships that included peer acceptance and peer rejection, two indexes that, although correlated, refer to different aspects of social life (Slaughter et al., 2015). This allowed us to identify specific patterns of associations. Finally, the use of a longitudinal design allowed us to investigate the effect of social relationships on pragmatic development and to shed light on the bidirectional nature of the associations between metaphor understanding and peer rejection.

Starting from these findings, future studies could also explore the causal nature of the observed longitudinal links. One way to address this issue would be to design an intervention program to promote metaphor understanding, such as the newly developed MetaCom training program (Tonini et al., 2021), and to examine its effect on children's social life. This could also have the practical implication of promoting a range of social functioning skills, thus reducing the risk of social exclusion by peers. We are not arguing that metaphor understanding per se might change children's social status, rather, we follow the idea that training conversational pragmatic skills has a positive impact on children's social life (Bierman & Furman, 1984; Ladd, 1981) and we propose that metaphor understanding should
be seen as the pinnacle of children's communicative capacities and, as such, its promotion might have a range of benefits for children's social relationships.

ACKNOWLEDGEMENTS
This work was supported by the MIUR (Ministero dell'Istruzione, dell'Università e della Ricerca) PRIN (Progetti di Ricerca di Rilevante Interesse Nazionale) 2015 project ‘The Interpretative Brain: Understanding and Promoting Pragmatic Abilities across Lifespan and in Mental Illness’, project code 201577HA9M, awarded to the last two authors.

CONFLICT OF INTEREST
None.

AUTHOR CONTRIBUTIONS
Valentina Bambini and Serena Lecce designed the research and the physical and mental metaphors task. Luca Ronchi collected the data. Paola Del Sette and Luca Ronchi coded the data and run the analyses. Paola Del Sette, Luca Ronchi, Valentina Bambini and Serena Lecce wrote the manuscript. All authors contributed to the final version of the manuscript with important intellectual content.

ENDNOTE
1 The composite reliability (CR) is computed using the following formula: $\frac{\left(\sum_{i=1}^{n} \lambda_i^2\right)}{\left(\sum_{i=1}^{n} \lambda_i^2 + \sum_{i=1}^{n} \text{VAR}(\epsilon_i)\right)}$ where $\lambda_i$ are the standardized factor loadings and $\text{VAR}(\epsilon_i)$ is the error variance.

DATA AVAILABILITY STATEMENT
The data that support the findings of the present study are available from the corresponding author upon request.

ORCID
Paola Del Sette https://orcid.org/0000-0001-6696-2139
Luca Ronchi https://orcid.org/0000-0001-7723-9376
Valentina Bambini https://orcid.org/0000-0001-5770-228X
Serena Lecce https://orcid.org/0000-0002-2745-3574

REFERENCES
Adamczyk, P., Daren, A., Sulecka, A., Bładziński, P., Cichocki, L., Kalisz, A., … Cechnicki, A. (2016). Do better communication skills promote sheltered employment in schizophrenia? Schizophrenia Research, 176(2–3), 331–339. https://doi.org/10.1016/j.schres.2016.08.015
Andersen, A., Krolner, R., Currie, C., Dallago, L., Due, P., Richter, M., … Holstein, B. E. (2008). High agreement on family affluence between children’s and parents’ reports: international study of 11-year-old children. Journal of Epidemiology & Community Health, 62(12), 1092–1094. https://doi.org/10.1136/jech.2007.065169
Andrei, F., Mancini, G., Mazzoni, E., Russo, P. M., & Baldaro, B. (2015). Social status and its link with personality dimensions, trait emotional intelligence, and scholastic achievement in children and early adolescents. Learning and Individual Differences, 42, 97–105 https://doi.org/10.1016/j.lindif.2015.07.01
Babad, E. (2001). On the conception and measurement of popularity: more facts and some straight conclusions. Social Psychology of Education, 5(1), 3–29. https://doi.org/10.1023/A:1012780232587
Babcock, B., Marks, P. E. L., Crick, N., & Cillessen, A. H. N. (2014). Limited nomination reliability using single- and multiple-item measures. Social Development, 23(3), 518–536. https://doi.org/10.1111/sode.12056
Bambini, V., Arcara, G., Bosinelli, F., Buonocore, M., Bechi, M., Cavallaro, R., & Bosia, M. (2020). A leopard cannot change its spots: A novel pragmatic account of concretism in schizophrenia. Neuropsychologia, 139, 107332. https://doi.org/10.1016/j.neuropsychologia.2020.107332
Banerjee, R., Watling, D., & Caputi, M. (2011). Peer relations and the understanding of faux pas: Longitudinal evidence for bidirectional associations. Child Development, 82(6), 1887–1905. https://doi.org/10.1111/j.1467-8624.2011.01669.x
Currie, C., Molcho, M., Boyce, W., Holstein, B., Torsheim, T., & Richter, M. (2008). Researching health inequalities in adolescents: the development of the Health Behaviour in School-Aged Children (HBSC) family affluence scale. *Social Science & Medicine, 66*(6), 1429–1436. https://doi.org/10.1016/j.socscimed.2007.11.024

Danneel, S., Colpin, H., Goossens, L., Engels, M., Leeuwen, K., Noortgate, W., & Verschueren, K. (2019). Emotional school engagement and global self-esteem in adolescents: Genetic susceptibility to peer acceptance and rejection. *Merrill-Palmer Quarterly, 65*(2), 158–182.

de Bruyn, E. H., & Cillessen, A. H. N. (2006). Popularity in early adolescence: Prosocial and antisocial subtypes. *Journal of Adolescent Research, 21*(6), 607–627. https://doi.org/10.1177/0743558406293966

Del Sette, P., Bambini, V., Bischetti, L., & Leccce, S. (2020). Longitudinal associations between theory of mind and metaphor understanding during middle childhood. *Cognitive Development, 56*, 100958. https://doi.org/10.1016/j.cogdev.2020.100958

Enders, C. K., & Bandalos, D. L. (2001). The relative performance of full information maximum likelihood estimation for missing data in structural equation models. *Structural Equation Modeling, 8*(3), 430–457. https://doi.org/10.1207/S15328007SEM0803_5

Filippova, E., & Astington, J. W. (2008). Further development in social reasoning revealed in discourse irony understanding. *Child Development, 79*(1), 126–138. http://doi.org/10.1111/j.1467-8624.2007.01151.x

Fite, P. J., Hendrickson, M., Rubens, S. L., Gabrielli, J., & Evans, S. (2013). The role of peer rejection in the link between reactive aggression and academic performance. *Child & Youth Care Forum, 42*(3), 193–205. https://doi.org/10.1007/s10566-013-9199-9

Flannery, K. M., & Smith, R. L. (2017). Are peer status, friendship quality, and friendship stability equivalent markers of social competence? *Adolescent Research Review, 2*(4), 331–340. https://doi.org/10.1007/s40894-016-0042-z

Fornell, C., & Larcker, D. F. (1981). Evaluating structural equation models with unobservable variables and measurement error. *Journal of Marketing Research, 18*(1), 39–50. https://doi.org/10.2307/3151312

Gao, Z., Gao, S., Xu, L., Zheng, X., Ma, X., Luo, L., & Kendrick, K. M. (2017). Women prefer men who use metaphorical language when paying compliments in a romantic context. *Scientific Reports, 7*, 40871. https://doi.org/10.1038/srep40871

Gazelle, H., & Rudolph, K. D. (2004). Moving toward and away from the world: social approach and avoidance trajectories in anxious solitary youth. *Child Development, 75*(3), 829–849. https://doi.org/10.1111/j.1467-8624.2004.00709.x

Gibbs, R. W. (2011). Evaluating conceptual metaphor theory. *Discourse Processes, 48*(8), 529–562. https://doi.org/10.1080/0163853X.2011.606103

Gifford-Smith, M. E., & Brownell, C. A. (2003). Childhood peer relationships: social acceptance, friendships, and peer net-

Glowacki, S. (1989). Metaphors in conversation: How are they understood? Why are they used? *Metaphor and Symbolic Activity, 4*(3), 125–143. https://doi.org/10.1207/s15327686ms0403_2

Glucksberg, S. (2003). The psycholinguistics of metaphor. *Trends in Cognitive Sciences, 7*(2), 92–96. https://doi.org/10.1016/S1366-6133(02)00040-2

Glucksberg, S., & Keysar, B. (1990). Understanding metaphorical comparisons: Beyond similarity. *Psychological Review, 97*(1), 3–18. https://doi.org/10.1037/0033-295X.97.1.3

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.

Hazen, N. L., & Black, B. (1989). Preschool peer communication skills: The role of social status and interaction context. *Child Development, 60*(4), 867–876. https://doi.org/10.2307/1131028

Horton, W. S. (2007). Metaphor and readers’ attributions of intimacy. *Memory & Cognition, 35*(1), 87–94. https://doi.org/10.3758/bf03195945

Hübischer, L., & Prieto, P. (2019). Gestural and prosodic development act as sister systems and jointly pave the way for children’s sociopragmatic development. *Frontiers in Psychology, 10*, 1259 https://doi.org/10.3389/fpsyg.2019.01259

Hughes, C. (2016). Theory of mind grows up: Reflections on new research on theory of mind in middle childhood and adolescence. *Journal of Experimental Child Psychology, 149*, 1–5. https://doi.org/10.1016/j.jexpchild.2016.01.017

Ji, L., Pan, B., Zhang, W., Zhang, L., Chen, L., & Deater-Deckard, K. (2019). Bidirectional Associations between peer relations and attention problems from 9 to 16 years. *Journal of Abnormal Child Psychology, 47*, 381–392. https://doi.org/10.1007/s10802-018-0440-8

Jiang, X. L., & Cillessen, A. H. N. (2005). Stability of continuous measures of sociometric status: A meta-analysis. *Developmental Mental Health Review, 25*(1), 1–25. https://doi.org/10.1016/j.dmr.2004.08.008

Kalandersd, T., Bambini, V., & Næss, K.-A. (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

Kintsch, W. (2001). Predication. *Cognitive Science, 25*(2), 173–202. https://doi.org/10.1207/s15516709cog2502_1
| Reference                                                                 | Details                                                                 |
|--------------------------------------------------------------------------|-------------------------------------------------------------------------|
| Serratrice, L. (2005).                                                  | The role of discourse pragmatics in the acquisition of subjects in Italian. |
| Selig, J. P., & Little, T. D. (2012).                                   | Autoregressive and cross-lagged panel analysis for longitudinal data.    |
| van der Ploeg, R., Kretschmer, T., Salmivalli, C., & Veenstra, R. (2017) | Defending victims: What does it take to intervene in bullying and how is it rewarded by peers? |
| van den Brink, D., Van Berkum, J. J. A., Bastiaansen, M. C. M., Tesink, C. | Efficacy and benefits of the MetaCom training to promote metaphor comprehension in typical development. (Manuscript submitted for publication). |
| Steen, G., Dorst, A., & Hermann, J. (2010).                             | Metaphor in usage. Cognitive Linguistics, 21, 765–796.                  |
| Thurstone, L. L., & Thurstone, T. G. (1962).                            | Primary mental abilities. Chicago, IL: Science Research.                |
| Tonini, E., Lecce, S., Del Sette, P., Bianco, F., Canal, P., & Bambini, V. | Efficacy and benefits of the MetaCom training to promote metaphor comprehension in typical development. (Manuscript submitted for publication). |
| Tudge, J. R., & Winterhoff, P. A. (1993).                               | Vygotsky, Piaget, and Bandura: Perspectives on the relations between the social world and cognitive development. Human Development, 36(2), 61–81. |
| van den Berg, Y. H. M., & Cillessen, A. H. N. (2012).                   | Computerized sociometric and peer assessment: An empirical and practical evaluation. International Journal of Behavioral Development, 37(1), 68–76. |
| van den Brink, D., Van Berkum, J. J. A., Bastiaansen, M. C. M., Tesink, C. | Empathy matters: ERP evidence for inter-individual differences in social language processing. Social Cognitive and Affective Neuroscience, 7(2), 173–183. |
| van der Plas, R., Kretschmer, T., Salmivalli, C., & Veenstra, R. (2017). | Defending victims: What does it take to intervene in bullying and how is it rewarded by peers? Journal of School Psychology, 65, 1–10. |
| van den Wilt, F., van der Veen, C., van Kruistum, C., & van Oers, B. (2018). | Why can’t I join? Peer rejection in early childhood education and the role of oral communicative competence. Contemporary Educational Psychology, 54, 247–254. |
APPENDIX A.

The internal consistency and the test–retest reliability of the PMM task were evaluated by conducting a series of confirmatory factor analyses (CFA). Specifically, we evaluated the internal consistency by assessing the dimensionality of the PMM latent structure and its scale reliability. We evaluated the test–retest reliability by assessing the metric invariance of the PMM latent structure and test–retest reliability between T1 and T2 (i.e., across a one-year period). These analyses were important to confirm that children’s response to each item of the PMM task did reflect a common underlying latent ability (i.e., metaphor accuracy) and that the latent structure of the PMM task was invariant and, therefore, reliable across time. In sum, these analyses were needed to examine whether the PMM task is a reliable and valid measure for assessing metaphor understanding in middle childhood.

The CFA model and the following analyses were performed on 11 items of the PMM, since a preliminary inspection of the proportions of children’s answers to each item of the PMM task revealed that the physical metaphor ‘Mio fratello è un grattacielo’ (literal translation ‘My brother is a skyscraper’ meaning that he is tall) was at ceiling, with 95% of children at T1 and 97% of children at T2 correctly expressing the salient feature of the metaphor topic. Thus, this item was excluded from all subsequent analyses. For the CFA models, given the categorical nature of our data, we used a mean- and variance-adjusted weighted least squares estimator (WLSMV; Brown, 2015; Kline, 2011). Nested CFA model comparisons were conducted using a corrected chi-square difference test suitable for WLSMV estimator. Through a CFA model in which all the items were regressed onto a single latent factor, we, therefore,
evaluated whether the 11 items of the PMM loaded onto a single latent factor. The resulting model was over-identified with $\chi^2(44) = 49.36, p = .268$, and showed a good fit to the data, RMSEA = 0.03, CFI = 0.94, TLI = 0.93. The standardized factor loadings of the PMM items on the latent factor ranged from 0.23 to 0.72 with a mean of 0.47 and all $p$s < .05. Altogether, these findings indicate that the PMM has a unidimensional factor structure. Crucially, the composite reliability (Fornell & Larcker, 1981) of the PMM factor was 0.76, indicating an adequate construct validity.

To test metric invariance of the PMM factor structure between T1 and T2, we evaluated a two latent factor CFA model in which PMM items at T1 loaded onto a single latent factor (i.e., T1 metaphor accuracy) and PMM items at T2 loaded onto a second latent factor (i.e., T2 metaphor accuracy), which was correlated to the T1 metaphor accuracy latent factor. Residual errors for each PMM item at T1 were correlated with residual errors for each corresponding PMM item at T2 in order to account for item-specific variance across time. The model provided a good fit to the data, $\chi^2(197) = 209.83, p = .252$, RMSEA = 0.02, CFI = 0.94, TLI = 0.93. Then, in order to test the equality of item loadings and latent factors variances between the two latent factors (i.e., metric invariance), the item loadings and latent factor variances were constrained to be equal across time points. A decrease in model fit, after the inclusion of these equality constraints, would indicate differences in the metric features of the metaphor accuracy latent factor across time. Results showed that the constrained model (Figure A1) continued to show good fit to the data, $\chi^2(208) = 218.51, p = .29$, RMSEA = 0.02, CFI = 0.95, TLI = 0.94, with no significant deterioration in model fit compared to the unconstrained model, $\Delta\chi^2(11) = 11.30, p = .42$. Given these results, we retained the constrained (more parsimonious) model, confirming the metric invariance for the metaphor accuracy latent factor across time. Importantly, given metric invariance, the metaphor accuracy latent factor showed strong test–retest reliability across 1 year ($r = .68, p < .001$).

![Figure A1](image-url)
In conclusion, CFA revealed that the PMM accuracy score has a unidimensional and reliable factor structure since both mental and physical items loaded onto a single underlying latent factor. The metaphor accuracy latent factor showed (a) adequate scale reliability, (b) metric invariance, and (c) strong test-retest reliability over 1 year. The unidimensionality of the PMM task is in line with experimental literature showing no differences in children’s ability to understand mental versus physical metaphors in middle childhood (see the accuracy measure in Lecce et al., 2019). Based on these results, for the analyses in this study, we computed a total metaphor accuracy score by summing the scores obtained in 11 metaphors (range 0–22).