UNIVERSITÀ
DI PAVIA

UNIVERSITA' DEGLI STUDI DI PAVIA

Dipartimento di Scienze del Sistema Nervoso e del Comportamento

DOTTORATO DI RICERCA IN
"Psychology, Neuroscience and Data Science"
Ciclo XXXIII

The social side of pragmatics in middle childhood: is metaphor understanding associated with Theory of Mind?

Tesi di
Paola Del Sette

SUPERVISORE: Prof.ssa Serena Lecce

COORDINATORE: Prof.ssa Gabriella Bottini

Anno Accademico 2019/2020

Trovarsi già in vetta sul crinale
davanti a nuovi cieli e nuove terre.

E, alle spalle, il gran coro delle voci
care e giovani, e il timbro di ciascuna:
il futuro degli altri che germoglia
da quell'unico tronco.

Sostare sentinella di confine
quasi in vista, e per ora solo a lampi,
di quel futuro libero dal tempo.

A. Capocaccia Quadri

Index

Abstract

Metaphor understanding is a pragmatic inferential skill that requires the ability to infer the speaker's intended meaning and serves a social function in adulthood. Given this definition, it seems reasonable that metaphor understanding skills are associated with Theory of Mind (ToM) abilities (i.e., the ability to infer others' mental states) and with social relationships. To date, the literature investigating these relationships is scarce. The present dissertation assesses this gap by investigating the associations between metaphor understanding and both social relationships and ToM in middle childhood, a crucial developmental period for both ToM and metaphor understanding. Moreover, middle childhood is a crucial period also for changes in social relationships that, in this period, become increasingly important for the development of children's socio-cognitive skills.

The present dissertation is made up of four studies.

The first study aimed to investigate the existence and the direction of the associations between metaphor understanding and peer relationships in middle-aged typically developing children. Preliminary, we were also interested in developing a tool for assessing metaphor understanding in middle childhood, namely the Physical and Mental Metaphors (PMM) task, and in assessing its psychometric properties. 126 9- to 10-year-old children were tested for their ability to understand metaphors (using an extended version of the PMM task) and for their peer relationships (assessing peer acceptance and rejection), at baseline and one year later. Results showed that (a) the PMM task is a sound measure for capturing individual differences in metaphor understanding during middle childhood, and (b) there is a longitudinal and bidirectional association between metaphor understanding and peer rejection, but

not peer acceptance. Through the first study, we conclude that being rejected hinders the development of metaphor comprehension and that understanding metaphors during conversations reduces peer rejection.

Starting from these results the *second study* investigated the existence of a link between metaphor understanding and ToM in typically developing middle-aged children. Preliminary, we were also interested in testing the psychometric properties of a short version of the PMM task, developed for assessing metaphor understanding. 217 children (aged 9 to 12 years) were assessed for their vocabulary, working memory, socio-economic status, ToM, and metaphors. In this second study, we decided to distinguish between metaphor comprehension and metaphor interpretation, an index that allowed us to make an important distinction between physical and mental metaphors. Preliminary results showed that the short version of the PMM task is an adequate measure for capturing individual differences in metaphor interpretation, but not in metaphor comprehension, during middle childhood. Main results showed that 9-year-olds (but not older children) who had higher ToM were also better in interpreting mental, but not physical metaphors. These findings lead us to conclude that the link between metaphorical skills and ToM is specific for the interpretation of mental, but not physical, metaphors and it is stronger in earlier rather than later childhood.

In the *third study,* we investigated the direction of the developmental relationships between ToM and metaphor understanding in middle-aged typically developing children using a short-term longitudinal design. In this study, we used the extended version of the PMM task that allowed us to distinguish between: (a) metaphor accuracy (the ability to find a link between metaphor topic and vehicle) and

specificity of mental interpretation (the ability to interpret mental but not physical metaphors, mentally), and (b) ToM and the ability to make inferences about physical states. 54 typically developing children (age-range = 8;6 – 9;4) were tested at baseline and 6 months later for inferential skills about mental (ToM) and physical states, and for metaphor understanding (accuracy and specificity of mental interpretation). Preliminary results on the soundness of the extended PMM task showed that it is a sound measure for capturing individual differences in metaphor comprehension and in specificity of mental interpretation, during middle childhood. In addition, the main results showed that Inferential skills about physical states were bi-directionally linked to the ability to comprehend a metaphor and that specificity of mental interpretation predicted later ToM (and not the reverse). We conclude that (a) the relationship between ToM and metaphor comprehension is mainly driven by general inferential abilities, (b) metaphor comprehension and inferential skills support each other during development, and (c) the tendency to mentally interpret mental, but not physical, metaphors drive ToM development.

Given the longitudinal link between some aspects of metaphor understanding and ToM, in the *fourth study,* we aimed to investigate the existence of causal relationships between ToM and metaphor understanding using a training design. We recruited 55 typically developing children (age-range = 8;7 – 10;3) and assigned them to two training conditions. Training conditions were made up of four sessions involving group conversations. In the ToM training, we trained children's ability to make context-sensitive inferences about mental states, while in the metaphor comprehension (MetaCom) training we trained children's ability to make context-sensitive inferences about metaphorical meanings. We measured, at baseline

and after the end of training programs ToM (via Strange Stories and animated triangles tasks) and metaphor understanding (via a hard version of the PMM and the referential tasks). Preliminary results on the soundness of the hard version of the PMM task showed that it is a measure able to capture individual differences in metaphor understanding, during middle childhood. Main results showed that while children in the ToM group improved in their ability to understand context-sensitive (but not minimal-context metaphors), children in the MetaCom group enhanced their ToM skills (assessed both with Strange Stories and animated triangles tasks). No training was effective in enhancing children's specificity of mental interpretation of metaphors. This pattern of results suggests that, while training ToM generalizes its effect on metaphorical comprehension by enhancing context-sensitive (but not minimal-context) inferences, training metaphor understanding (that includes reasoning about mental interpretation) improves children's ability to attribute and infer others' mental states.

In conclusion, the findings from these four studies give support to the existence of a link between metaphor understanding and both social relationships and ToM. In addition, regarding the associations between metaphor understanding and ToM, they supported a view in which their relationships are not fixed but vary depending on the specific pragmatic aspect considered. Thus, while ToM predicts and causes the development of metaphor comprehension, working on the general ability to make inferences about contextual and phrasal information, the tendency to refer to mental states for mental, but not physical, metaphors drives and causes ToM development.

From a more general point of view, our findings provide support to the claim that ToM and pragmatics do not overlap.

General introduction

Pragmatics is a complex theoretical construct that has been defined in several different ways throughout years (Bosco, Tirassa, & Gabbatore, 2018). Behind such different perspectives, there is agreement that pragmatic ability requires to go beyond the literal meaning of an utterance using inferential processes to fill the gap between what is explicitly said and what is communicated (Arcara & Bambini, 2016; Searle, 1979; Sperber & Wilson, 2005). These inferential processes, that enable individuals to understand speakers' intended meaning and to interpret incoming utterances, derive from contextual information (Sperber & Wilson, 2002; Zufferey, 2010). Given this definition, pragmatics would be necessary to navigate the social world and for individuals' daily social exchanges (Bohn & Frank, 2019). In line with this, developmental studies showed a parallelism in the impairment of both social functioning and communicative pragmatic skills in children with Autism Spectrum Disorder (ASD) (American Psychiatric Association, 2013; Philofsky, Fidler, & Hepburn, 2007), Williams syndrome (Laws & Bishop, 2004; Philofsky et al., 2007) and specific language impairment (Conti-Ramsden & Botting, 2004; St Clair, Pickles, Durkin, & Conti-Ramsden, 2011). Accordingly, in ASD children, deficits in pragmatic skills were also found to predict children's poor social functioning (Volden, Coolican, Garon, White, & Bryson, 2009) and to mediate the relationship between ASD symptoms and social relationships (Berenguer, Miranda, Colomer, Baixauli, & Roselló, 2018). Moreover, research conducted on typically developing samples have shown that individual differences in pragmatics were associated with variation in

children's social position within the peer group during the pre-school years (see for a review, van der Wilt, van der Veen, van Kruistum, & van Oers, 2019).

Nevertheless, despite both theoretical and empirical works support the importance of pragmatic abilities for children's social life and development, a clear understanding of the nature of the association between pragmatics and children's social relationships is far from being accomplished (van der Wilt, et al., 2019). In detail, little is known about whether and how individual differences in pragmatics influence and/or are influenced by children's social relationships. This is even more interesting when children grow older and start to spend more and more time with peers in the school setting (Rubin, Bukowski, & Parker, 2007).

Starting from this, and given that individual differences in children's peer relationships are related to their ability to understand others' mental states (i.e., Theory of Mind [ToM]) in pre-school and school years (Banerjee, Watling, & Caputi, 2011; Caputi, Lecce, Pagnin, & Banerjee, 2012; Hughes, 1998; Slaughter, Imuta, Peterson, & Henry, 2015; Watson, Nixon, Wilson, & Capage, 1999), it seems reasonable that pragmatics is related also with ToM. Accordingly, several studies found a relationship between pragmatics and ToM in children with typical and atypical development (see for a review, Bosco et al., 2018). However, little is known about the relationship between ToM and the specific pragmatic function of metaphor understanding. In addition, it is still not clear if individual differences in pragmatics influence ToM and/or are influenced by ToM.

The present dissertation fits within this very recent field of research to investigate, during middle childhood, the developmental interplay between

children's pragmatics, their peer relationships, and their ToM, with a specific focus on metaphor understanding.

The development of pragmatics

How children come to infer the intended meaning of an utterance has attracted the attention of developmental researchers coming from two different traditions: the one that linked the emergence of communication to language acquisition and the one that used the linguistic pragmatics framework (Bohn & Frank, 2019). The literature on the origins of language described children's communicative abilities from the first year of age, while studies on the development of linguistic pragmatics focused on children at ages of 3 to 6 years (Bohn & Frank, 2019). Taking together the evidence coming from the two different frameworks, the development of pragmatic inference seems to have its origin in children's first years and continues to develop throughout their lives (Bohn & Frank, 2019).

Several studies in the developmental domain showed that children by 6–9 months of age have already all the ingredients of mature pragmatic inference (Bohn & Frank, 2019). Indeed, human communication makes use of pragmatic inference early in age, for example through the so-called ostensive signals (i.e., eye contact or infant-directed speech). All those signals lead infants and children to interpret others' actions as communicative (Bohn & Frank, 2019). For example, from 5-7 months of age infants are more likely to follow an actor's gaze to a target when it is preceded by gaze or speech cues (Hernik & Broesch, 2019; Senju

& Csibra, 2008). However, 6-month-old infants can follow gaze even in non-ostensive contexts (Gredebäck, Astor, & Fawcett, 2018; Szufnarowska, Rohlfing, Fawcett, & Gredebäck, 2014). Ostensive cues are, therefore, sufficient but not necessary for communicative inference. As children grow, they start to use pragmatic inferences in support of language learning. For example, some studies showed that from 17 months, children infer that novel words refer to novel objects (Frank & Goodman, 2014; Halberda, 2003; Markman & Wachtel, 1988; Markman, Wasow, & Hansen, 2003).

By the age of 2 years, children, in order to reduce the cost of producing unnecessary utterances, can already select what to say based on what the partner already knows because of a common ground (Bohn & Frank, 2019). For example, they produce more information (naming objects and producing referential gestures) about a hidden object and when the social partner is unaware of the hiding (O'Neill, 1996). In the same vein, 2-year-old children produce more linguistic information when the gesture of pointing would be ambiguous alone (O'Neill & Topolovec, 2001). By the age of 3 to 5 years, children start to select the kind of information given, depending on whether their social partner knows the facts and the objects he/she speaks about (Baer & Friedman, 2018; Köymen, Mammen, & Tomasello, 2016). In addition, children from 3 years of age start to select not only the specific information to include in the utterance, but also which words to use based on partner-specific expectations about referential expressions (Bohn & Frank, 2019; Matthews, Lieven, & Tomasello, 2010). After the age of 3 years, children start to have enough vocabulary and grammatical ability to

interpret more complex utterances and to infer word meanings from the linguistic context (Gleitman, 1990). In line with this hypothesis, 4- to 5-year-old children are faster to fixate on an object when a familiar speaker refers to this object with a previously used expression compared to a new speaker using the same expression (Graham, Sedivy, & Khu, 2013).

In the field of pragmatic inferences during conversations, implicatures are a core aspect. Implicatures seem to be a little more difficult compared to the pragmatic inferences we have already discussed. Indeed, 5-year-old and even older children still struggle to compute implicatures, especially under time constraints (Huang & Snedeker, 2009; Noveck, Bianco, & Castry, 2001). This may depend on the fact that during implicatures in communication children have to generate alternatives that could not be easily available (Bohn & Frank, 2019). This interpretation suggests that children struggle, not necessarily with making pragmatic inferences per se, but with generating the alternatives that are necessary for making inferences (Barner, Brooks, & Bale, 2011; Bohn & Frank, 2019).

Given the definition according to which lexical pragmatics require to go beyond the literal meanings of words to interpret language in context, high order pragmatic skills are required. Thus, the inferences required to understand non-literal language, such as idioms, metaphors, and irony, are more complex than the ones described above. Indeed, the ability to understand idioms, ironic statements, and metaphors starts to develop from the age of 5-6 years and continues to improve throughout development. In detail, *idioms* are not understood until the age of 6 (Abkarian, Jones, & West, 1992) and their metapragmatic knowledge

appears after the age of 9 years (Bernicot, Laval, & Chaminaud, 2007). The comprehension of *irony* starts around the age of 6 years (Dews & Winner, 1997; Harris & Pexman, 2003), or sometimes even earlier (Angeleri & Airenti, 2014; Loukusa & Leinonen, 2008), but continues to develop even after the age of 13 years. For example, a study found that 9-year-old children do not reach adults' ability to understand ironic statements (Filippova & Astington, 2008). Accordingly, it has been shown that children as old as 13 years could fail to distinguish irony from deception (Demorest, Meyer, Phelps, Gardner, & Winner, 1984; Demorest, Silberstein, Gardner, & Winner, 1983; Winner, Levy, Kaplan, & Rosenblatt, 1988). Similarly, research on metaphorical skills has clearly shown that *metaphor understanding* undergoes a developmental progression over school years. The process of metaphor understanding (e.g., "Dancers are butterflies"), requires children to attribute to the topic (i.e., "Dancers") the most salient property of the vehicle (i.e., "butterflies are light") conveying the metaphorical meaning (i.e., "Dancers are light") (Wilson & Carston, 2007). Especially when tested with verbal explanation tasks, young children are unable to articulate the link between the topic and the vehicle (Winner, Rosenstiel, & Gardner, 1976). Indeed, they start to develop metaphor understanding from the age of 6 years, even if, at this age, some children continue to interpret metaphorical utterances as literally true (e.g., "Dancers fly"). At this developmental stage, children make more errors by inferring a relationship between the topic and the vehicle based on association rather than similarity (e.g. "Dancers are dressed like butterflies") (Winner et al., 1988). From 8 to 10 years of age, children typically make errors in

understanding metaphors basing their pragmatic inference about a non-salient similarity between the topic and the vehicle (e.g. “Dancers are small”) (Winner et al., 1988). In conclusion, children’s ability to understand metaphors undergoes a developmental progression over school years with a key turning point at around 9 years of age (Noveck et al., 2001). For this reason, focusing on middle childhood would be crucial when investigating developmental gains and individual differences in children’s metaphor understanding.

The development of peer relationships

Peer interactions represent an extremely important aspect of children’s lives, having a great influence on children’s development of social, emotional, and cognitive functioning (e.g., Banerjee et al., 2011; Lecce, Bianco, & Ronchi, 2020). Interestingly, peer relationships undergo several changes during children’s development.

From the first year of life, infants’ peer interactions, which at that age are diffuse and fragmented, undergo remarkable strikes (Eckerman, 1979; Hay, Nash, & Pedersen, 1983; Rubin et al., 2007). From the second year of age, with the emergence of locomotion and language, toddlers demonstrate monumental gains in their social repertoires and start to organize their social interactions around games. During toddlerhood children begin to (a) coordinate their behavior with the one of the social partners, (b) imitate peer’s behavior, (c) appropriately communicate with peers, basing communicative responses to the partners’

characteristics, and (d) help, share and be in conflict with others (Rubin et al., 2007).

Throughout the pre-school years, children increase the time spent interacting with peers: they start to spend time simply conversating with their playmates about their numerous interpersonal goals (Hay, Payne, & Chadwick, 2004; Rubin et al., 2007). From 2 to 5 years of age, peer interactions become more frequent and more complex. Indeed, by the 3rd year, children start to share with others symbolic meanings through pretend play (Howes, 1988; Rubin et al., 2007). During pretense, children practice their ability to share meaning with their social partner and, during pre-school years, become better in understanding terms of shared meaning (Göncü, 1993; Rubin et al., 2007). Pretend play serves several developmental functions since it provides a safe context for mastering communication, compromise, negotiation, and for exploring others' perspectives (Göncü, 1993; Howes, Unger, & Matheson, 1992; Sawyer, 1997). Thus, before school years, children acquire an important milestone for their social lives: the understanding of a pretending scenario and the ability to share this pretending scenario with others (Rubin et al., 2007). In conclusion, all these changes in children's social lives during pre-school years, lead to an improvement in their social-communicative competence (Dunn, 1999; Krasnor & Rubin, 1983; Shatz & Gelman, 1973).

As children enter primary school, their social relationships change dramatically in terms of amount, importance, and complexity (Hughes, 2016). In terms of the number of social relationships, the proportion of peer interactions

increases, becoming more than 30% of the total amount of social interactions (Rubin et al., 2007). In terms of importance, during middle childhood, children's social interactions become increasingly demanding and reliant on children's own cognitive and social skills rather than on adults' support (Rubin et al., 2007). In terms of complexity, during school years, children start to come into contact with peers in several different settings, including conversing, "hanging out", being together at school, talking on the telephone, traveling to and from school, listening to TV and records, and playing non-contact sports (Rubin et al., 2007; Zarbatany, Hartmann, & Rankin, 1990). In addition, the complexity of social relationships during middle childhood is also reflected by the new settings in which social interactions take place. From the primary school years, the structure of the peer group changes from a relatively unified whole to more differentiated structures in which children organize themselves into social groups (Bagwell, Coie, Terry, & Lochman, 2000; Değirmencioğlu, Urberg, Tolson, & Richard, 1998). The primary organizational feature of middle-aged children's groups is the popularity hierarchy that could be measured via the sociometric and the perceived popularity (Coie, Dodge, & Coppotelli, 1982). Through the sociometric procedure, researchers can obtain a measurement of *acceptance,* that reflects the extent to which a child is liked by her/his peers, and *rejection,* that mirrors the extent to which a child is actively rejected by her/his peers. *Perceived popularity* refers to a child's perceived status in the peer group and, therefore, differs from acceptance in meaning and measurement (Hawley, Little, & Pasupathi, 2002; Lease, Kennedy, & Axelrod, 2002).

Within the peer social context children continue to improve in their social, emotional, and cognitive functioning acquiring a wide range of behaviors, skills, and experiences that influence their future adaptation. Indeed, from middle childhood, experiences with peers start to represent a risk factor, affecting their subsequent academic and psychological adjustment (Rubin et al., 2007).

The development of Theory of Mind

Theory of Mind (ToM) is defined as the ability to understand mental states in order to predict and interpret others' behavior (Wimmer & Perner, 1983). How children learn to use ToM has been largely investigated by developmental research in the last 40 years.

During children's development, ToM undergoes significant improvement (Wellman & Liu, 2004). In the first year of life, children start to take into account others by imitating their facial expressions and by pointing objects to request something (imperative pointing) and, later on, to draw others' attention to an interesting object (declarative pointing). Imperative and declarative pointings are identified as "precursors" of ToM development.

After the first year of life, children start to become familiar with others' mind: they start to use a mental lexicon, for referring to self and others' mental states, to appreciate that others have desires and beliefs that can differ from their own, and to understand that something can be true but someone might not know that (Wellman & Liu, 2004).

Between the ages of 3 to 5 years, children face a core developmental shift in their ability to use mental states (beliefs, intentions, knowledge, desires) to explain others' behaviors. Indeed, they start to understand that others' behaviors are driven by beliefs that can also be false. Accordingly, between 3 and 5 years of age, typically developing children usually pass the false-belief task[1] (Wellman, Cross, & Watson, 2001). The ability to pass 1st order false-belief tasks is considered a milestone of children's ToM development. Indeed, understanding a false belief was considered as a marker of children's ability to understand the meta-representative nature of the mind. Thus, passing a false belief task requires children the ability to adapt the cognitive perspective of someone else in terms of mental states and to desist from reality as well as from one's representation (Wimmer & Perner, 1983).

Recently, studies on ToM development extend into middle childhood showing that middle-aged children become progressively better in understanding others' false beliefs and, more generally, others' mental states (Hughes, 2016; Perner & Wimmer, 1985). Indeed, studies, adopting advanced ToM tasks, revealed important progression in children's ToM during middle childhood (Banerjee, Watling, & Caputi, 2011; Lecce, Bianco, Devine, & Hughes, 2017; Ronchi, Banerjee, & Lecce, 2020). Some authors focused on an *interpretative*

[1] The false-belief task assesses children's ability to recognize that an agent has a false belief about the location or the content of an object. During the task, children are asked to predict agent's behavior based on his/her false belief.

ToM, that is the ability to understand that a single situation can be open to multiple possible interpretations (Carpendale & Chandler, 1996). Studies focusing on this kind of task revealed that children started to develop the interpretative ToM by the age of 6 to 7, and continued to grow throughout middle childhood (Lagattuta, Sayfan, & Blattman, 2010; Lagattuta, Sayfan, & Harvey, 2014). Other authors focused on the ability to understand how mental states can affect people's behaviors within complex social scenarios, for example through the Strange Stories and the Silent Films tasks (Devine & Hughes, 2013; Devine, White, Ensor, & Hughes, 2016; Happé, 1994). These kinds of tasks are able to grasp developmental and individual differences during middle childhood, since they measure children's ability to *use* previously acquired mental states knowledge in a more flexible and appropriate way, to make sense to their social environment (Hughes & Devine, 2015). Accordingly, middle-aged children become increasingly better at considering and integrating several different mental states with contextual information to make flexible and appropriate predictions about others' intentions and behaviors (Hughes, 2016). For example, Lecce and colleagues (2017) reported a significant improvement in children's ToM abilities, assessed via the Strange Stories and the Silent Films tasks, over a period of 6 months in 9-year-old children. In addition, even older children were found to continue improving. For example, a recent study found that 11-year-olds improved in their ToM, assessed via the Strange Stories and the Silent Films tasks, over a year (Ronchi et al., 2020). For these reasons, focusing on developmental gains and individual differences in children's ToM during middle

childhood is crucial both from a theoretical and an applied perspective (Hughes, 2016).

Pragmatics and peer relationships

Studies examining the association between social relationships and basic pragmatic conversational skills (e.g., the ability to contribute to the communication with relevant comments and the ability to regulate turn-taking) found relationships between pre-school children's pragmatic conversational abilities and their level of acceptance and/or rejection within the peer group (Black & Hazen, 1990; Black & Logan, 1995; Hazen & Black, 1989; van der Wilt et al., 2019; van der Wilt, van der Veen, van Kruistum, & van Oers, 2020; van der Wilt, van Kruistum, van der Veen, & van Oers, 2016). The few studies focused on middle childhood found controversial results: some studies found a link between children's ability to conversate effectively and their social relationships (Leonard, Milich, & Lorch, 2011; Miranda, Berenguer, Roselló, & Baixauli, 2020; Petranovich, Walz, Staat, Chiu, & Wade, 2017; Putallaz & Gottman, 1981; Putallaz & Wasserman, 1989; Rabiner & Gordon, 1992), while others did not find any relationship (Bierman & Furman, 1984; Nowicki & Oxenford, 1989). The presence of controversial results in middle childhood could be explained by the fact that pragmatic conversational skills are low-level pragmatic abilities that could be already mastered before the beginning of school (Blain-Brière, Bouchard, & Bigras, 2014; Hübscher & Prieto, 2019; Serratrice, 2005). Thus,

during school years children would show little variation in these more basic conversational competencies (van der Wilt et al., 2019).

In the previous chapters, we have highlighted middle childhood as a crucial developmental period for high-level pragmatics skills such as those required to understand non-literal language (e.g., idioms, metaphors, and irony). In addition, we have also pointed out that social relationships become more complex during middle childhood. Thank to this complexity, peer relationships become really important for the development of cognitive skills, including social understanding (Banerjee et al., 2011), executive functions (Lecce et al., 2020), cognitive engagement, and academic performances (Buhs, & Ladd, 2001; Chen, Hughes, Liew, & Kwok, 2010; Fite, Hendrickson, Rubens, Gabrielli, & Evans, 2013; Oberle & Schonert-Reichl, 2013; Wentzel & Caldwell, 1997). Indeed, interpersonal interactions, including exchanges with peers, act in the zone of children's proximal development (Vygotsky, Cole, John-Steiner, Scribner, & Souberman, 1978), allowing them to learn, practice, and co-construct cognitive skills and knowledge (Rubin et al., 2007).

Given this parallelism in developmental timing, it should be interesting to investigate the association between high-level pragmatic skills and social relationships in middle childhood. Focusing on a specific high-level pragmatic skill, namely metaphor understanding, no study has directly investigated the association between metaphorical skills and social relationships in middle childhood. This lack of empirical studies is very striking, given that, during adulthood, metaphors were found to have a social function. At least three sources

of evidence bring support to the social function of metaphors. First, studies showed that metaphors make the social communication more persuasive (Landau, Meier, & Keefer, 2010; Sopory & Dillard, 2002) and the interlocutor more attractive (Gao et al., 2017). Second, metaphors were found to increase participants' sensitivity to social cues during conversations by creating intimacy between the listener and the speaker (Bowes & Katz, 2015; Cohen, 1978; Horton, 2007). Indeed, the listener and the speaker of a metaphorical utterance are drawn closer to one another through the creation of common ground and of shared knowledge (Cohen, 1978). Third, clinical studies found that the impairment of metaphor comprehension hinders individual's social functioning and relationships (Adamczyk et al. 2016; Bambini et al. 2020). In addition, indirect empirical support to the idea that metaphor understanding is associated with children's peer relationships come from the studies described above that examined the association between social relationships and other aspects of pragmatic abilities in pre-schoolers and school-age children (see for a review, van der Wilt et al., 2019).

Pragmatics and ToM in middle childhood

In the previous chapters, we have highlighted that Theory of Mind and high-level pragmatic inferential skills underwent a parallel development, with middle childhood as a key developmental period. Given this parallelism, research focused on middle childhood when investigating the association between ToM and high-level pragmatic inferential skills. For example, several studies focused on the relationships between the understanding of irony and ToM. One of the

main reasons that prompt the researchers to focus on this association is the process underlying irony understanding. Indeed, the listener, in order to comprehend an ironic statement, should understand the speaker's attitude in producing such a statement (Matthews, Biney, & Abbot-Smith, 2018). The literature on this topic confirms the expected association by showing, in school-age children, significant relationships between ToM and irony comprehension (e.g., Filippova & Astington, 2008; Nilsen Glenwright, and Huyder 2011).

Regarding the interplay between ToM and metaphor understanding, the literature is more scant. The vast majority of this literature focused on the clinical population supporting the existence of this association (Champagne-Lavau & Stip, 2010; Happé, 1993; Huang, Oi, & Taguchi, 2015; Kalandadze, Norbury, Nærland, & Næss, 2016; Mo, Su, Chan, & Liu, 2008; Norbury, 2005). Till now, only one study explored this relationship also on typically developing children (Whyte & Nelson, 2015). Results showed that, in typically developing children, aged from 5 to 12 years, nonliteral language performance, including (but not limited to) metaphor comprehension skills, was related to ToM, even after controlling for basic language abilities. This study was not exempt from limitations, since it considered a wide age range and was not focused on metaphor understanding skills.

To date, no other study further investigated the association between metaphor understanding and both peer relationships and ToM in typically developing children during primary school years.

The present dissertation

The present dissertation fits within this developmental literature and adopts a developmental approach to investigate (a) the existence and the direction of the association between pragmatics and peer relationships, (b) the existence of an association between pragmatics and ToM, (c) the direction of the association between pragmatics and ToM, and (d) the causal nature of the association between pragmatics and ToM. In order to reach these goals, the present dissertation is made up of 4 studies. Each study addresses one of the above questions. In each study, a focused approach was adopted making important two choices: (1) we concentrated on a key developmental stage for the constructs under study, that is, middle childhood; (2) we assessed a single specific ability of the more general pragmatic skills, namely metaphor understanding. In doing so, the preliminary aim of the present dissertation was to develop a tool for assessing metaphor understanding in middle childhood, namely the Physical and Mental Metaphors (PMM) task. Thus, while investigating the main aim of each study, we also tested the psychometrical properties of this task trying to create the best and the shortest version for assessing metaphor understanding, being sensitive to individual differences in middle childhood.

Regarding the structure of the dissertation, each chapter describes an experimental study built to deal with one major question.

The *first chapter* describes a longitudinal study built to investigate the existence and the direction of the associations between metaphor understanding and peer relationships in middle childhood. In addition to this main aim, Study 1

aimed to examine the psychometric properties of an extended version of the PMM task.

The *second chapter* describes a cross-sectional study aiming to investigate the existence of the association between metaphor understanding and ToM in middle childhood. Preliminarily, Study 2 aimed to test the psychometric properties of a short version of the PMM task and to investigate the developmental changes of metaphor understanding and ToM in children from 9 to 12 years.

The *third chapter* describes a short-term longitudinal study investigating the direction of the associations between metaphor understanding and ToM in middle childhood. In addition to the main aim, we were also interested in examining (a) the developmental trajectories of ToM and metaphor understanding, (b) the rank-order stabilities of the key variables over 6 months, and (c) the concurrent relationships between metaphor understanding and ToM skills.

Finally, the *fourth chapter* describes a training study investigating the causal relationships between ToM and metaphor understanding during middle childhood.

In each chapter, it is reported a detailed rationale, the specific hypotheses, measures, and the method adopted in each study. See Appendix A for a summary of all the measures included in each study of the present dissertation.

Chapter 1 – Metaphor Understanding and social relationships: longitudinal associations in middle childhood

Abstract

The aim of the present longitudinal study was to investigate the existence and the direction of the associations between metaphor understanding and peer relationships in middle-aged typically developing children. Preliminary, we were also interested in developing and testing the psychometric properties of a tool for assessing metaphor understanding in middle childhood, namely the Physical and Mental Metaphors (PMM) task. We assessed, in 126 typically developing 9-year-old children, their ability to understand metaphors (via the PMM task) and their peer relationships (assessing peer acceptance and peer rejection through the nomination procedure), at baseline and one year later. Results showed that the PMM task is an adequate measure for capturing individual differences in metaphor understanding during middle childhood. In addition, we found a longitudinal and bidirectional association between metaphor understanding and peer rejection, but not peer acceptance. We conclude that being rejected hinders the development of metaphor comprehension and that not understanding metaphors during conversations increases peer rejection.

Introduction

Metaphor understanding is an advanced pragmatic skill that requires the ability to fill the gap between the literal and the intended meaning of a sentence

conveyed by a social partner (Carston, 2012; Sperber & Wilson, 2005). To be understood metaphors require the ability to go beyond the literal meaning in order to understand the speaker's intended meaning (Sperber & Wilson, 2005). Nominal metaphors, such as "Climbers are squirrels", are typically composed of a topic, that is, the subject of the metaphor ("Climbers"), and a vehicle, that is, the term used to convey the metaphorical meaning ("squirrels"). During the process of understanding, individuals infer the intended topic's features through the literal meaning of the vehicle (Dent & Rosenberg, 1990; Wilson & Carston, 2007). Usually, communicative exchanges that involve metaphors require the activation of shared social knowledge and higher sensitivity to social cues (Bowes & Katz, 2015). For this reason, metaphors are important during social exchanges and create intimacy by drawing the listener and the speaker closer to one another (Bowes & Katz, 2015; Cohen, 1978; Horton, 2007). Thus, metaphors serve a social function that leads them to be commonly used in everyday conversations during adulthood (Cameron, 2008; Glucksberg, 1989) and sometimes to be preferred to literal sentences (Ortony, 1975; Sopory & Dillard, 2002).

The social function of metaphors has so far been studied on adults, with the consequence that very little is known about the social role of metaphor understanding in children. This lack is striking especially for middle childhood, considering that this age, as anticipated in the general introduction, is a crucial time for both social relationships and metaphor understanding. On one hand, middle childhood is a key period for studying the development of social skills. Indeed, children's social life changes dramatically as they enter primary school

since their social relationships start to become increasingly demanding, to mainly involve peers, and to be less supervised by adults (Rubin et al., 2007).

On the other hand, studies have clearly shown that metaphor understanding undergoes a developmental progression over school years, with a key turning point during middle childhood (Noveck et al., 2001; Winner et al., 1976).

To conclude, given that middle childhood is a key period for both social relationships and metaphor understanding, focusing on this developmental stage would be crucial when studying the interplay between peer relationships and metaphor comprehension.

Empirical support to the idea that metaphor understanding is related to children's peer relationships comes from studies examining the association between social relationships and other aspects of pragmatic abilities, namely children's pragmatic conversational skills (e.g., the ability to contribute to the communication with relevant comments and the ability to regulate turn-taking). However, even if these studies found relationships between children's pragmatic conversational abilities and their level of acceptance and/or rejection within the peer group they mainly focused on pre-schoolers (Black & Hazen, 1990; Black & Logan, 1995; Hazen & Black, 1989; van der Wilt et al., 2019; van der Wilt, van der Veen, van Kruistum, & van Oers, 2020; van der Wilt, van Kruistum, van der Veen, & van Oers, 2016). In addition, the few studies focused on middle childhood found controversial results. In detail, while some studies found a link between children's ability to conversate effectively and (a) their social position

within the peer group, (b) their social skills problems, and (c) their social competence, assessed via parents' reports (Leonard, Milich, & Lorch, 2011; Miranda, Berenguer, Roselló, & Baixauli, 2020; Petranovich, Walz, Staat, Chiu, & Wade, 2017; Putallaz & Gottman, 1981; Putallaz & Wasserman, 1989; Rabiner & Gordon, 1992), others did not find any relationship (Bierman & Furman, 1984; Nowicki & Oxenford, 1989). This absence of effects in middle childhood could be in part be determined by the fact that pragmatic conversational skills are low-level pragmatic abilities that develop rapidly throughout early childhood (Bohn & Frank, 2019; van der Wilt et al., 2019) and that could be already in place for most children during primary school years. In line with this, most of the studies that found a link between social relationships and pragmatics in middle childhood are those that have used more comprehensive assessments of both pragmatics (such as the pragmatic scale of the Children's Communication Checklist, CCC; Bishop, 1998, 2013) and social skills (such as Strength and Difficulties Questionnaire, SDQ; Goodman, 1997) (Leonard et al. 2011; Miranda et al., 2020; Petranovich et al., 2017). Even if these studies are relevant, as they suggest that those children who have more sophisticated pragmatic abilities are indeed more socially competent compared to those who lag beyond in pragmatic development, they still have limitations. Perhaps, one of the most important limits is that these studies have measured pragmatics and social skills using parents' reports and this may have inflated the correlations between the two.

Summarizing the limits of the literature described above, till now studies investigating the genuine association between individual differences in pragmatics

and the one in peer relationships focused on a low-level pragmatic ability or used parents' reports for measuring social skills and pragmatics. Therefore, future studies should focus on (a) pragmatic measures that are able to capture individual differences in middle childhood, like high order pragmatic inferential skills, and (b) direct measures of both constructs. The present study tried to follow these recommendations for the assessment of both peer relationships and pragmatic skills.

Regarding peer relationships, the present study adopted direct measures assessing both peer acceptance and peer rejection through the nomination procedure. Considering these two aspects of social relationships (i.e., peer acceptance and peer rejection) is crucial, since, although moderately associated, they are different from one another and have different correlates (Slaughter et al., 2015). Indeed, peer acceptance refers to the number of most-liked nominations received by classmates and reflects how much a child is liked within the class, while peer rejection refers to the number of least-liked nominations received by classmates and reflects how much a child is explicitly disliked by peers.

Regarding pragmatics, the present study adopted a direct measure of a high order pragmatic ability, specifically metaphor understanding. In the literature on middle childhood, the ability to comprehend metaphors is usually investigated through tasks that require children to paraphrase sentences with a nominal metaphor and thus to make inference about the topic meaning within a phrasal context (Pinto, Melogno, & Iliceto, 2006, 2008; Winner, Engel, & Gardner, 1980; Winner et al., 1976). Indeed, these kinds of tasks are able to capture children's

individual differences during school years (Winner et al., 1976), while simpler tasks, such as elicited repetition or picture selection tasks, reveal rudimentary metaphorical abilities already in pre-school children (Pouscoulous, 2014).

Interestingly, some authors emphasized differences related to the types of metaphors, highlighting the importance to distinguish between cross-sensory metaphors (communicating an experience in one sensory modality by referring to another sensory modality; e.g., "Her perfume was bright sunshine") and psychological-physical metaphors (communicating a psychological experience by referring to the physical domain; e.g., "The prison guard was a hard rock") (Melogno, Pinto, & Di Filippo, 2017; Nippold, Leonard, & Kai, 1984; Winner et al., 1976). Till now, there are controversial types of evidence about the different difficulties to understand those two metaphors. While some authors suggested that the cross-sensory metaphors were easy to be understood compared to psychological-physical metaphors (Vosniadou, Ortony, Reynolds, & Wilson, 1984; Wang & Dowker, 2010; Winner et al., 1976), others found that they were equally difficult to be understood (Nippold et al., 1984). Despite interesting, till now, the literature that considers this difference is scarce. Using tasks that allow to make the distinction between physical (i.e., cross-sensory) and mental (i.e., psychological-physical) metaphors would help to better understand the links between specific pragmatic skills and other aspects, for instance, social relationships.

The present study expanded existing literature also by adopting a longitudinal design which, allowed us to address not only the existence of

relationships but also the direction of the expected associations between metaphor understanding and social relationships. Indeed, all the studies described above were cross-sectional and for this reason, little is still known about the directionality of the associations between peer relationships and pragmatics in typically developing middle-aged children. The few studies adopting longitudinal/experimental design provided evidence on the direction according to which children's sociometric status could in part be determined by their conversational pragmatic skills during interactions with peers (Gifford-Smith & Brownell, 2003; Newcomb, Bukowski, & Pattee, 1993). In detail, evidence, suggesting that communicative skills predict later social relationships, comes from longitudinal (Burleson, Delia, & Applegate, 1992; Law, Rush, Clegg, Peters, & Roulstone, 2015) and training studies (Bierman & Furman, 1984; Ladd, 1981). While the hypothesis according to which pragmatic skills affect social competencies was largely investigated by empirical studies, no study explored empirically the opposite hypothesis (i.e., that social interaction could influence the development of pragmatic skills). However, the latter hypothesis is still reasonable according to the socio-constructivists theoretical models suggesting that the development of cognition relies on social interactions (Vygotsky et al., 1978). In line with these models, a recent position paper suggested that the emergence of metaphorical use could have its origins in social interaction with peers (Clark, 2019). Indeed, playing with peers (e.g., pretend play) could help children to develop the ability to re-categorize the role of objects (e.g., a pen

becomes a sward), and that this, in turn, allows children to extend the use of conventional terms beyond the literal meaning of words (Clark, 2019).

The Present Study

The main aim of the present study was to investigate the existence and the direction of the developmental associations between pragmatic inferential skills, specifically metaphor understanding, and social relationships, as indexed by peer acceptance and peer rejection. Interestingly, we adopted a newly developed task for measuring metaphor understanding, namely the Physical and Mental Metaphors (PMM) task. The PMM task distinguished between physical vs. mental metaphors, two kinds of metaphors that could be differently related to social relationships. Preliminarily, we were therefore interested in examining the factorial structure and the test-retest reliability of this newly developed task.

Given the social function of metaphors in adulthood (Bowes & Katz, 2015; Cohen, 1978; Horton, 2007) and the burst of both social relationships and metaphor understanding in middle childhood (Rubin et al., 2007; Winner et al., 1976), our main hypothesis was the existence of longitudinal and bidirectional associations between social relationships and metaphor comprehension in middle childhood. In detail, given existing empirical evidence showing that conversational skills influence the development of positive social relationships (Bierman & Furman, 1984; Burleson et al., 1992; Ladd, 1981; Law et al., 2015; Place & Becker, 1991), we expected that this effect extended to high-level pragmatics inferential skills, like metaphor understanding. In addition, we

expected to find also an effect in the opposite direction (i.e., that social relationships predict the development of metaphor understanding) based on socio-constructivists theoretical models and a recent position paper claiming that having positive social interactions and playing with peers help children to develop cognitive skills, especially metaphorical skills (Clark, 2019; Rubin et al., 2007; Vygotsky et al., 1978). More precisely, we hypothesized that children with a better understanding of metaphors became less rejected and more accepted and that children who were more accepted and less rejected were more likely to develop a better understanding of metaphors.

To test this bidirectional hypothesis, we designed a two-wave cross-lagged longitudinal study in which we measured individual differences in children's metaphor understanding (through the Physical and Mental Metaphors task) and social relationships (via children's peer nomination) at baseline and one year later.

Given the literature suggesting that receptive language could have a role in the relationships between social acceptance/rejection and pragmatics (van der Wilt et al., 2019, 2020), we controlled for this variable in our main analyses. Moreover, given that van der Wilt and colleagues (2016) found an association between peer relationships and pragmatics only in boys and not in girls, we also considered the moderator role of gender in the expected longitudinal associations between metaphor understanding and peer relationships.

Materials and Method

Participants

One hundred sixty-one Italian children, attending Year 4 of the primary school participated in the present study. Among these, 33 were not eligible due to specific learning disorders, developmental delays or not being native Italian speakers. The final sample consisted of 126 typically developing school-aged children (59 M), aged 9-10 years at Time 1 (*M*age = 9 years; 10 months, *SD* = 0;3, age-range = 9;2–10;6). All children who took part *at least* at one time-point were included in our analyses. The percentage of missing values ranged from 1.56% to 8.59% (see Table 1.2 for detailed sample sizes). The rate of missing data was highest at Time 1 metaphor understanding due to some new children recruited at Time 2. In the modeling analyses, we handled missing data by the full information maximum likelihood estimation (FIML; Enders & Bandalos, 2001).

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 (Time 1, May) and one year later (Time 2, May). At Time 1 children were also assessed for vocabulary abilities and were asked to complete a family affluence questionnaire. This study has been approved by the local University Ethical Committee.

Measures

Socio-economic status. Socio-economic status was assessed at Time 1 through the Family Affluence Scale (FAS; Currie et al., 2008). In this questionnaire, children were asked to answer four questions evaluating the family ownership of a car (range = 0–2), the ownership of an unshared bedroom (range = 0–1), the number of computers at home (range = 0–3), and the number of family vacations in a 12-month period (range = 0–3). An overall index of the family's socio-economic background was obtained by summing the score at each question (range = 0–9).

Vocabulary. Vocabulary was measured at Time 1 using the Italian version of the vocabulary subtest of the Primary Mental Abilities, intermediate form (age 11-17) (PMA; 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. The final score was given by the number of correct responses (range = 0–50).

Peer relationships. Children indices of peer acceptance and peer rejection were obtained through the sociometric peer nomination procedure (Coie et al., 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 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 Participants section) 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 has shown to have high test-retest reliability (Jiang & Cillessen, 2005). Moreover, indices of peer acceptance and peer rejection show construct validity since variability in the number of positive (ML) and negative (LL) nominations by peers was found to be highly associated with social behavior (Rubin et al., 2007) and socio-cognitive understanding (Banerjee et al., 2011). Concerning internal consistency, previous studies consistently found reliability scores around .70 for peer rejection (Babcock, Marks, Crick, & Cillessen, 2014; Marks, Babcock, Cillessen, & Crick, 2013; van den Berg & Cillessen, 2012) and around.10 to .40 for peer acceptance (Babad, 2001; Babcock et al., 2014). In order to evaluate the internal consistency score for peer acceptance and peer rejection, in our sample, we used the Kuder–Richardson Formula 20 (Kuder & Richardson, 1937), following the procedure described in Cillessen and Marks (2017). In line with previous literature cited above, our mean

reliability scores across classrooms were .71 (*SD* = .13) at Time 1 and .79 (*SD* = .06) at Time 2 for peer rejection, and .28 (*SD* = .31) at Time 1 and .41 (*SD* = .39) at Time 2 for peer acceptance. Finally, test-retest reliability across 1 year, calculated through both correlations and path analyses, was high for both peer acceptance and peer rejection ($r_s > .70$, $p_s < .001$).

Metaphor understanding. We constructed a Physical and Mental Metaphors (PMM) task composed of a set of 14 nominal metaphors in Italian, following the prototypical "X is Y" structure, where X is the topic and Y the vehicle of the metaphor, and distinguishing between two types of metaphors: physical and mental. Persons (X) are always associated to non-human entities (Y). In physical metaphors, the associations were based on physical features while in mental metaphors were based on psychological features. Thus, in order to understand physical metaphors, inferences about appearance features or behaviors of the topic are needed, whereas, in order to understand mental metaphors, inferences about mental states of the topic are needed. Physical metaphors were: (1) "Il cuoco è una botte" (literal translation "The chef is a barrel"), (2) "Il calciatore è una freccia" ("The footballer is an arrow"), (3) "Mio fratello è un grattacielo" ("My brother is a skyscraper"), (4) "le ballerine sono farfalle" ("Dancers are butterflies"), (5) "Gli scalatori sono scoiattoli" ("Climbers are squirrels"), (6) "I giocatori sono elefanti" ("Players are elephants"), and (7) "Quel giornalista è un fiume" ("That journalist is a river"). Mental metaphors were: (8) "Quell'alunno è una spugna" (literal translation "That pupil is a sponge"), (9) "La

nonna è una colonna" ("Grandma is a column"), (10) "La mamma è un cioccolatino" ("Mummy is a candy"), (11) "I soldati sono Leoni" ("Soldiers are lions"), (12) "Il papa è un vulcano" ("Daddy is a volcano"), (13) "La maestra è un ghiacciolo" ("The teacher is an icicle") and (14) "Quell'uomo è un orso" ("That man is a bear").

The only difference between the two sets is the preferred interpretation (either referred to physical or mental attributes of the topic of each metaphor) since we selected the metaphors balanced for lexical frequency and familiarity. Several data were collected to verify that those characteristics are respected. For a full description of PMM items see Table 1.1a. First, we checked frequency including all target words in a frequency dictionary of child language (Marconi, Ott, Pesenti, Ratti, & Tavella, 1994). Even if the frequency of target words showed a wide range, we found no difference between the physical and mental sets, $t(12) = .80$, $p = .442$.

Second, we administered an online rating task in which 53 young adults were asked to rate familiarity (i.e., the grade of familiarity of each metaphor), and aptness (i.e., the degree to which a metaphor vehicle captures appropriate and relevant features of a metaphor topic), using a 7-point Likert-type scale. Thus, we compared familiarity and aptness of mental and physical metaphors. Results on familiarity indicated that all metaphors (seven mental and seven physical) were non-lexicalized (i.e., no metaphor was highly familiar), with no differences across sets, $t(12) = -1.71$, $p = .114$. Results on aptness showed all metaphors (seven mental and seven physical) were apt (i.e., no metaphor was low in aptness) and

that metaphor in the mental set did not differ from metaphor in the physical set in the grade of appropriateness in which the metaphor describes the salient characteristics of the metaphor topic, *t(12)* = -1.08, *p* = .300.

Third, the distinction between physical and mental metaphors was checked through two tasks: (a) the parametric task, and (b) the categorical task.

(a) In the parametric task 53 young adults blind to the hypothesis of the study were asked to rate each metaphor for physical attributes ("How much does this metaphor express physical attributes?") and for mental attributes ("How much does this metaphor express mental attributes?") using a 7-point Likert-type scale. As expected, the seven metaphors in the physical set scored significantly higher on physical attributes compared to the seven metaphors in the mental set, $t(12) = 3.15$, $p = .008$. On the contrary, the seven metaphors in the mental set scored significantly higher on mental attributes compared to the seven metaphors in the physical set, $t(7.59) = -6.60$, $p < .001$. Moreover, within each set, the difference between the two interpretations (the absolute value of the difference between the mean score on the physical attributes and the mean score on the mental attributes) was significantly bigger for the physical set compared to the mental set, $t(12) = 3.76$, $p = .003$.

Table 1.1a. Characteristics of the items of the extended version of the Physical and Mental Metaphor task.

Metaphor	English literal translation	Set	Excluded Metaphor	Familiarity	Physical score	Mental score	Aptness	Frequency
Il cuoco è una botte	The chef is a barrel	Physical	No	4.08	5.64	1.44	4.52	16.12
Il calciatore è una freccia	The footballer is an arrow	Physical	No	3.88	5.88	1.56	4.88	64.95
Mio fratello è un grattacielo	My brother is a skyscraper	Physical	No	4.44	6.04	1.60	4.76	6.50
Le ballerine sono farfalle	Dancers are butterflies	Physical	No	5.07	6.18	1.57	5.46	96.23
Gli scalatori sono scoiattoli	Climbers are squirrels	Physical	No	3.11	5.68	1.43	4.64	75.39
I giocatori sono elefanti	Players are elephants	Physical	No	2.76	4.53	1.92	2.80	85.88
Quel giornalista è un fiume	That journalist is a river	Physical	Yes	2.50	2.43	4.64	3.64	276.62
Quell'alunno è una spugna	That pupil is a sponge	Mental	No	5.25	2.46	5.68	5.18	15.23
La nonna è una colonna	Grandma is a column	Mental	No	3.92	2.76	5.32	4.56	27.98
La mamma è un cioccolatino	Mummy is a candy	Mental	No	4.20	3.00	5.08	4.56	13.96
I soldati sono leoni	Soldiers are lions	Mental	No	4.32	4.29	5.04	4.89	158.83
Il papa è un vulcano	Daddy is a volcano	Mental	No	3.64	3.25	4.71	4.07	24.83
La maestra è un ghiacciolo	The teacher is an icicle	Mental	No	4.16	3.16	5.24	4.68	5.10
Quell'uomo è un orso	That man is a bear	Mental	Yes	5.93	4.57	4.54	5.71	142.12

Notes: In the table we report the ratings of familiarity, physical and mental characteristics, aptness and frequency values in school-age children language (based on Marconi et al., 1994). These ratings were obtained from 53 subjects (40 F; *M*age = 23.91, *SD* = 2.33, age-range = 21 – 32; *M*education = 15.83, *SD* = 1.67, education range = 13 – 18).

(b) In the categorical task, 32 young adults blind to the hypothesis of the study were asked to classify each metaphor as being referred either to physical or mental attributes. Results led us to the exclusion of two metaphors from the task. In detail, the metaphor "That man is a bear" was not easily attributable to one of the two sets as there was no difference in the two attribution preferences, $\chi^2(1) = 0.50$, $p = .480$. In addition, the metaphor "That journalist is a river" was interpreted, contrary to our expectation, as conveying a mental, rather than a physical, meaning, $\chi^2(1) = 4.50$ $p = .034$. Thus, the final extended version of the PMM task was composed of six mental and six physical metaphors.

We decided to repeat the previous analyses on the 12-items final version of the PMM task and found results that were very similar to the ones reported for the 14-items version for frequency, familiarity, aptness, and parametric task (Table 1.1b). In addition, the 12-items version overcame the limits of the 14-items version showing, via the categorical task, that metaphors in the physical set were considered physical in more than 91% and that metaphors in the mental set were considered mental in more than 72%. For each metaphor, we checked if the frequency observed was different from a casual distribution, finding highly significant chi-squared for each set (see Table 1.1a) and for each metaphor, $\chi^2(1)_s > 6.12$ $p_s < .014$. Overall, these results bring us to two main conclusions. First, they supported the distinction between the physical and the mental set. Second, they indicated that metaphors in the physical set tended to have a straightforwardly physical interpretation, while metaphors in the mental set were also open to physical interpretations.

Table 1.1b. Comparison between the mental set and the physical set of the extended version of the Physical and Mental Metaphor task.

		All materials	Physical set	Mental set	Statistics
Frequency		49.25 (47.49)	57.51 (37.40)	40.99 (58.31)	$t(10) = 0.58$, $p = .572$
Familiarity		4.07 (0.71)	3.89 (0.85)	4.25 (0.55)	$t(10) = -0.87$, $p = .405$
Aptness		4.58 (0.66)	4.52 (0.90)	4.66 (0.37)	$t(10) = -0.37$, $p = .721$
Categorical Task	Agreement on physical interpretation	-	97.40% [91-100]	2.60% [0-9]	$X^2(1) = 172.52$, $p < .001$
	Agreement on psychological interpretation	-	10.42% [3-28]	89.58% [72-97]	$X^2(1) = 120.33$, $p < .001$
Parametric Task	Physical attributes	4.41 (1.43)	5.66 (0.59)	3.15 (0.62)	$t(10) = 7.14$, $p < .001$
	Mental attributes	3.38 (1.89)	1.59 (0.18)	5.18 (0.32)	$t(10) = -23.89$, $p < .001$
	Difference score	-	4.07 (0.73)	2.03 (0.85)	$t(10) = 4.45$, $p = .001$

Notes. For frequency, familiarity, aptness and the parametric task scores, the Table reports mean and SD in parenthesis. These ratings were obtained from 53 subjects (40 F; *M*age = 23.91, *SD* = 2.33, age-range = 21 – 32; *M*education = 15.83, *SD* = 1.67, education range = 13 – 18). For the categorical task scores, the Table reports mean agreement and range in parenthesis. These ratings were obtained from 32 subjects (22 F; *M*age = 22.35, *SD* = 1.99, age-range = 21 – 30; *M*education = 15.16, *SD* = 0.95, education range = 14 – 20).

During the task, after an example item that was considered together with the examiner, children were required to explain orally the meaning conveyed by each of the 12 metaphors**.** Children's answers were coded according to the level of accuracy defined as the ability to articulate the link between the topic and the vehicle. Accuracy score is based on the assumption that the meaning of the nominal metaphor can be fully grasped in the light of salience. Indeed, the most accurate understanding of metaphors is described as the one that is conceptually salient, in the light of conventionality, frequency, familiarity, and prototypicality (Giora, 2003). Following this definition, accuracy was coded on a three-point scale and scores for each response ranged between 0 and 2. A 0 score was assigned for 'don't know' answers, incorrect answers (e.g., for the metaphor "Il papa è un volcano" [Daddy is a volcano], "È buono e dolce" [he is nice and sweet]), and literal answers ("Erutta" [He erupts]). A score of 1 was assigned for answers that were incomplete or referred to non-salient features of the metaphor topic ("È grande" [He is really big] "È caldo" [He is hot]). A score of 2 was assigned when the children's answer was complete and referred to the salient features of the metaphor vehicle ("Si arrabbia" [He gets angry]).

A second rater independently coded 25% of the responses and inter-rater agreement was established using Cohen's kappa, showing an almost perfect agreement at both time points (k = .91 at Time 1 and k = .88 at Time 2). A total accuracy (range = 0–24) was calculated by summing the score obtained at each metaphor (see Preliminary analyses for the psychometric properties of this measure).

Statistical analyses plan

Before testing our main hypothesis (i.e., the existence of bidirectional associations between metaphor understanding and peer relationships over time), we evaluated the psychometric properties of the PMM task and conducted preliminary analyses. Concerning the psychometric properties of the PMM task, we adopted a Confirmatory Factor Analyses (CFA) framework and evaluated: (a) the dimensionality of the PMM latent structure and scale reliability at baseline, and (b) the metric invariance and the test-retest reliability of the PMM latent structure across 1 year (i.e., between Time 1 and Time 2).

Regarding preliminary analyses, we presented descriptive statistics for all the study measures and conducted a mixed ANCOVA to examine developmental changes in metaphor accuracy, as well as gender differences within and across time. We also calculated the correlation matrix for all the study variables.

To verify our main hypothesis (i.e., the existence of reciprocal associations between metaphor understanding and peer relationships), we adopted an autoregressive cross-lagged modeling approach (see Selig & Little, 2012) controlling for language, family affluence, and longitudinal stabilities. Finally, to test for possible differences in the hypothesized longitudinal associations between metaphor accuracy and peer relationships across males and females, we used a multiple-group modeling approach with gender as the grouping variable.

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). For the path analysis 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 (χ^2) test, a root mean square error of approximation (RMSEA) $\leq .08$, a comparative fit index (CFI), and a Tucker-Lewis index (TLI.) $\geq .90$. Nested CFA model comparisons were conducted using a corrected chi-square difference test suitable for WLSMV estimator. Differently, a Satorra-Bentler chi-square difference test (Satorra, 2000; Satorra & Bentler, 2001) suitable for MLR estimator was used to compare nested path-analysis models.

Results

Psychometric properties of the Physical and Mental Metaphors task

A preliminary inspection of children's answers to each item of the PMM task revealed that the metaphor "My brother is a skyscraper" was at ceiling, with 95% of children at Time 1 and 97% of children at Time 2 referring to the salient feature of the topic. We, therefore, excluded this item from all subsequent analyses.

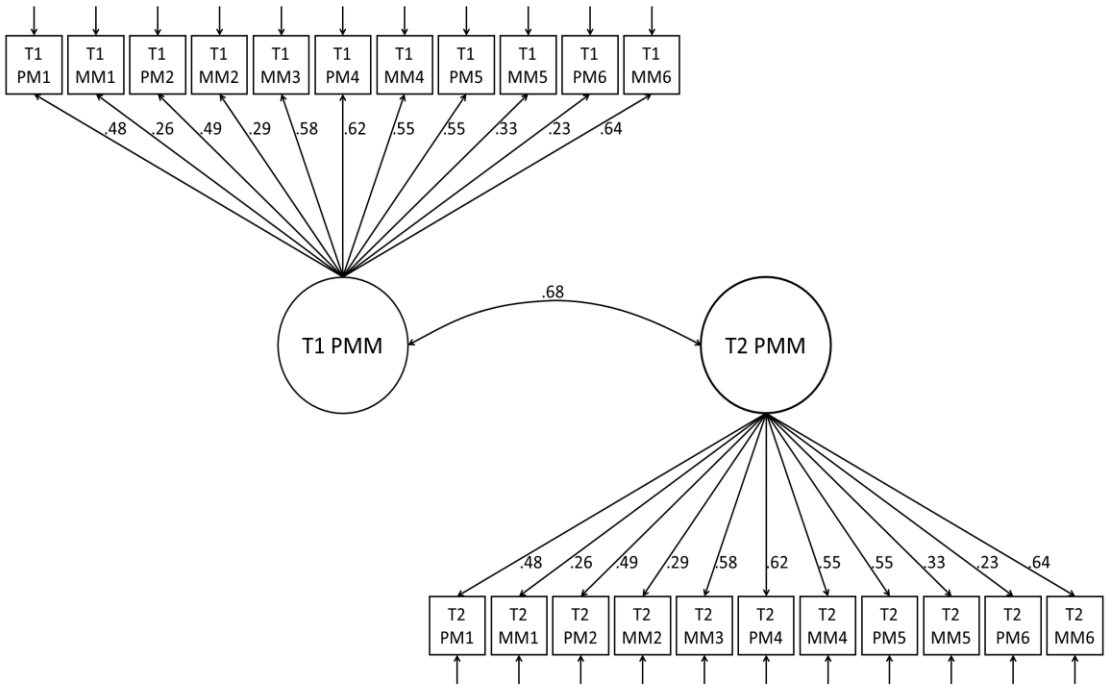
To test if, as expected, physical and mental metaphors were separated constructs, we tested a two latent factors model for the PMM task. Thus, we applied a parallel categorical data CFA, testing a model in which the accuracy score of the five physical items of the PMM task loaded on a Physical Metaphors latent factor (PM), and the six mental metaphor loaded on a Mental Metaphor (MM) latent factor. The resulting model was over-identified with, $\chi^2(43) = 48.84$,

p = .250, and showed a good model fit, *RMSEA* = .03, *CFI* = .94, *TLI* = .92. The standardized factor loadings of the two-factor model ranged from .24 to .73 with a mean of .48 and all p_s < .05. In order to verify the bi-dimensionality of PMM we then compared this model to a more parsimonious one-factor model in which all 11 items of the PMM task loaded on a single latent factor. The resulting model continued to provide good fit to the data $\chi^2(44) = 49.36$, p = .268, *RMSEA* = .03, *CFI* = .94, *TLI* = .93 with no significant deterioration in model fit compared to two-factor model $\Delta\chi^2(1) = 0.40$, p = .526, suggesting that the two-factor solution was not superior to the one-factor model. Altogether these findings suggested that, contrary to our expectations, the accuracy score of the PMM task had a unidimensional factor structure. The standardized factor loadings of the final one-factor model ranged from .23 to .72 with a mean of .47 and all p_s < .05. Given the unidimensional factor structure of the accuracy score, we performed subsequent analyses using a total accuracy score.

We then tested the invariance of the unifactorial structure of the PMM task between Time 1 and Time 2; we estimated a two latent factor measurement model in which all the PMM items at Time 1 loaded onto a single latent factor (i.e., Time 1 metaphor accuracy) and all the PMM items at Time 2 loaded onto a second correlated latent factor (i.e., Time 2 metaphor accuracy). In order to account for item-specific variance across time, residual terms for each PMM item at Time 1 were correlated with residual terms for each corresponding PMM item at Time 2. The model provided a good fit to the data $\chi^2(197) = 209.83$, p = .252, *RMSEA* = .02, *CFI* = .94, *TLI* = .93. To test metric invariance of this model, item loadings

and latent factor variances were constrained to be equal across time points and we evaluated if there would be a decrease in model fit after the inclusion of equality constraints. The second (more parsimonious) constrained model (Figure 1.1) continued to provide a good model fit, $\chi^2(208) = 218.51$, $p = .29$, $RMSEA = .02$, $CFI = .95$, $TLI = .94$, with no significant deterioration compared to the unconstrained model $\Delta\chi^2(11) = 11.30$, $p = .42$. Thus, our results confirmed the metric invariance for the PMM latent factor across time. Notably, the PMM latent factor showed strong test-retest reliability across 1-year period ($r = .68$, $p < .001$).

Figure 1.1. Standardized estimates for the final measurement model for the Physical and Mental Metaphors task with item loadings and latent factor variances constrained to be equal across time points.



Notes. PM = Physical Metaphor; MM = Mental Metaphor; PMM = Physical and Mental Metaphors task; T1 = Time 1; T2 = Time 2. All parameter estimates were $p < .018$.

Preliminary analyses

Descriptive statistics of all study variables are reported in Table 1.2.

Table 1.2. Descriptive Statistics.

Tasks	Valid N	Mean (SD)	Actual Range
T1			
Family Affluence Scale	120	2.72 (0.47)	1–3
Vocabulary	121	18.69 (5.51)	6–34
Most-Like nomination	121	0.15 (1.00)	-1.62–3.16
Least-Like nomination	121	-0.04 (0.95)	-1.31–3.89
Metaphor Accuracy	115	15.14 (3.81)	0–21
T2			
Most-Like nomination	124	0.12 (1.02)	-1.59–3.60
Least-Like nomination	124	-0.10 (0.93)	-1.02–3.35
Metaphor Accuracy	123	16.72 (3.41)	0–22

Notes. Metaphor Accuracy is reported after excluding the ceiling item; therefore, the possible range is 0–24. T1 = Time 1; T2 = Time 2.

We performed a mixed ANCOVA on metaphor accuracy with time as a within-subject variable and gender as a between-subject variable. We found a main effect of time, $F(1,110) = 19.79$, $p < .001$, $\eta_p^2 = .15$, but no gender effect, $F(1,110) = .0003$, $p = .99$, $\eta_p^2 < .001$, and no time by gender interaction effect, $F(1,110) = .005$, $p = .945$, $\eta_p^2 < .001$. We, therefore, found that metaphor accuracy improved from Time 1 to Time 2, while there were no differences in metaphor accuracy between male and female. These results did not change even after controlling for vocabulary, $F(1,109) = 5.76$, $p = .018$, $\eta_p^2 < .05$, $F(1,109) = .08$, $p = .78$, $\eta_p^2 = .001$, $F(1,109) = .0002$, $p = .988$, $\eta_p^2 < .001$, respectively.

Table 1.3 reports bivariate correlations among all the study measures. Each of the key variables showed moderate to high rank-order stability across time ($r_s > .45$, $p_s < .001$). Regarding the within-time correlations, there was a

significant negative correlation, at Time 2, but not at Time 1, between metaphor accuracy and LL. Differently, ML was not related at any time point to metaphor accuracy.

Table 1.3. Bivariate and Partial Correlations (below the diagonal) among study variables.

	T1			T2		
Tasks	ML	LL	MAcc	ML	LL	MAcc
T1						
FAS	.11	.25**	.07	.14	-.18+	-16
Vocabulary	.03	.02	.32**	.12	-.10	.24*
ML	—	-.42***	-.03	.47***	-.36***	-.036
LL	-.40***	—	-.12	-.25**	.66***	-.18+
MAcc	-.07	-.12	—	-.06	-.25**	.45***
T2						
ML	.46***	-.26**	-.11	—	-.39***	.05
LL	-.33***	.68***	-.22*	-.37***	—	-.27**
MAcc	-.04	-.19*	.40***	.03	-.27**	—

Notes. Partial correlations controlled for Vocabulary and Family Affluence. T1 = Time 1; T2 = Time 2; FAS = Family Affluence Scale; ML = Most Like nominations; LL = Least Like nominations, MAcc = Metaphor Accuracy. + $p <$ .06. * $p < .05$. ** $p < .01$. *** $p < .001$.

Longitudinal associations between peer relationships and metaphor understanding

In order to investigate reciprocal longitudinal associations between peer relationships (indexed by ML and LL nominations) and metaphor understanding (indexed by metaphor accuracy), 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

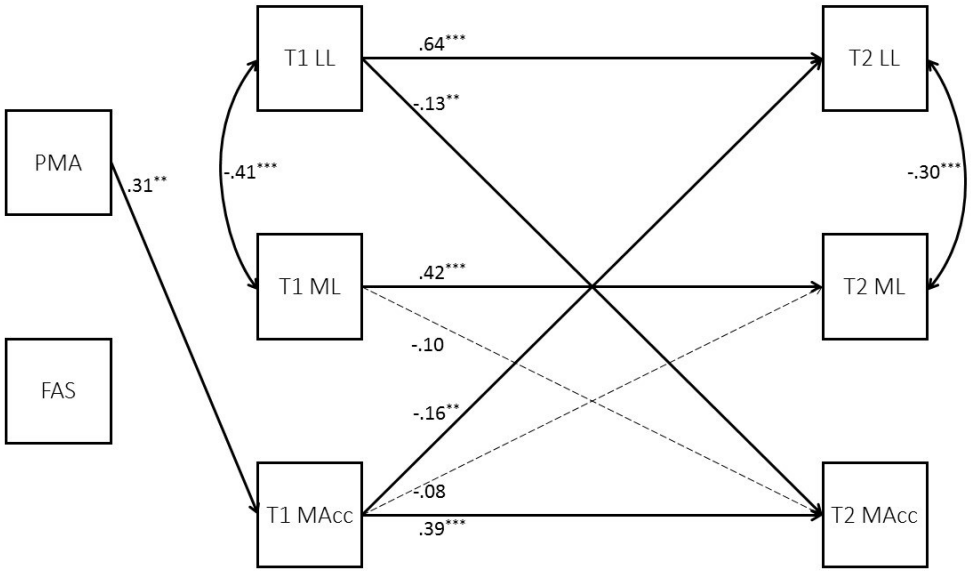
paths linking (a) metaphor accuracy at Time 1 with ML and LL at Time 2, and (b) ML and LL at Time 1 with metaphor accuracy at Time 2. In the model we also included family affluence and vocabulary as control variables, predicting all the variables at both Time 1 and Time 2. Finally, we permitted error 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, $\chi^2(2) = 2.969$, $p = .23$, $RMSEA = .06$, $CFI = .99$, $TLI = .92$.

Results showed that early LL nominations exerted a significant negative effect on later metaphor accuracy, $B = -0.18$, $\beta = -0.17$, $p = .039$ and that early metaphor accuracy exerted a significant negative effect on later LL nominations, $B = -0.12$, $\beta = -0.14$, $p = .046$. On the contrary, ML nominations did not show any significant association with metaphor accuracy across time (in both directions).

Given the reciprocal and longitudinal nature of the association between LL nominations and metaphor accuracy, we further tested whether the strength of this association was equal across directions. Thus, we tested a model in which the cross-lagged paths between LL nominations and metaphor accuracy were constrained to be equal in both directions and evaluated a possible decrease in the model fit after the equality constrain inclusion. The constrained model continued to provide a good fit to the data, $\chi^2(3) = 3.42$, $p = .33$, $RMSEA = .03$, $CFI = 1.00$, $TLI = .98$, and, compared to the previous unconstrained model, showed no statistically significant decrease in the model fit, $\Delta\chi^2(1) = .33$, $p = .565$. This pattern of results indicated that the two regression paths (i.e., from early metaphor accuracy to later LL nominations and from early LL nominations to later

metaphor accuracy) were not statistically different in strength. Figure 1.2 shows the final constrained model with standardized parameter estimates.

Figure 1.2. Path diagram depicting longitudinal relations between social relationships, measured via least liked and most liked nominations and metaphor understanding.



Notes. Paths with $p > .05$ are shown in dotted line. FAS = score in the Family Affluence Scale; PMA = score in the vocabulary subtest of the Primary Mental Ability; T1 LL = Least Liked nomination at Time 1; T1 ML = Most Liked nomination at Time 1; T1 MAcc = Metaphor Accuracy at Time 1; T2 LL = Least Liked nomination at Time 2; T2 ML = Most Liked nomination at Time 2; T2 MAcc = Metaphor Accuracy at Time 2. ** $p < .01$. *** $p < .001$.

Finally, in order to test for possible differences across gender in the hypothesized longitudinal paths, we adopted a multi-group procedure comparing a model in which cross-lagged paths were freely estimated across males and females to a nested model in which all cross-lagged paths were constrained to be equal across gender. We found no significant decrease in the model fit, suggesting

that there were no differences between males and females in the strength of the association between LL nominations and metaphor accuracy in any direction, $\Delta\chi^2(4) = 6.20$, $p = .185$.

Discussion

The present study aimed to test the existence and the direction of associations between metaphor understanding and social relationships in middle-aged children.

Before discussing the main results, it is important to highlight our preliminary findings concerning the psychometric properties of the PMM task. The PMM task was designed to take into account the distinction between physical and mental aspects of metaphors. Some studies have suggested an existing difference across metaphor types (Melogno et al., 2017; Nippold et al., 1984; Winner et al., 1976). Contrary to those studies, our results suggest that children's ability to understand physical metaphors and their ability to understand mental metaphors are not separate constructs. Indeed, we showed that the PMM accuracy score had a unidimensional factor structure. This evidence is in line with the body of experimental literature that found no differences in the ability to understand mental vs. physical metaphors (Nippold et al., 1984). Interestingly, regarding the developmental trajectories of metaphor understanding, we found that 9- to 10-year-olds improved in their ability to comprehend a metaphor over a year even over and above individual differences in vocabulary skills. This result fits with previous evidence suggesting that metaphor comprehension starts to be accurate at

about 10 years of age (Winner et al., 1976). Notably, this result suggests that the PMM task is a sound measure for capturing individual differences in metaphor understanding during middle childhood. Finally, other evidence of the PMM task psychometrical soundness comes from our findings showing that the task exhibited metric invariance between time points and good scale and test-retest reliability. Altogether these results indicate that the task has satisfactory statistical properties for measuring metaphor understanding in middle childhood.

Moving to the main goal, the present study aimed to investigate, in the critical period of middle childhood, the existence and the direction of the links between a high level inferential pragmatic skill, namely metaphor understanding, and social relationships. In detail, we wanted to explore the developmental associations between metaphor understanding and two distinct indices of social relationships, namely peer rejection and peer acceptance. Following empirical evidence and theoretical models presented in the introduction, we hypothesized to find bidirectional relationships between metaphor understanding and both peer rejection and peer acceptance indices. Interestingly, our analyses showed that the association between metaphor understanding and peer relationships exists in both directions over time for peer rejection but not for peer acceptance. Notably, we found that this developmental association was equal for females and males and was independent of individual differences in vocabulary, family affluence, and longitudinal stability in key variables.

The unexpected findings showing that only peer rejection, but not peer acceptance, was longitudinally and bi-directionally related to metaphor

understanding can be explained by the fact that peer acceptance and peer rejection are not simply opposites but, rather, reflect two distinct dimensions of children's social life (Rubin et al., 2007). Indeed, children who score high on peer acceptance are not only those who are popular (having high acceptance and low rejection) but also those who are controversial (having high acceptance and also high rejection). In the same way, children who score low on peer acceptance are not only those who are rejected (having low acceptance and high rejection) but also those who are neglected (having low acceptance and low rejection). In line with the view according to which peer rejection and peer acceptance are different constructs, empirical findings showed that they had different correlates (Rubin et al., 2007; Slaughter et al., 2015) and reflected separate dimensions of social experience (Gifford-Smith & Brownell, 2003; Lecce, et al., 2020). Indeed, while peer acceptance was found to be related to children's sensitivity, prosociality, and emotions regulation (Chen et al., 2010; Gifford-Smith & Brownell, 2003), peer rejection was found to be associated with negative peer treatment, externalizing problem behaviors, and decline in classroom participation (Bierman, 2004; Buhs, & Ladd, 2001; Kraatz-Keiley, Bates, Dodge, & Pettit, 2000; Ladd & Troop-Gordon, 2003).

Regarding the main aim (i.e., the existence and direction of the relationships between metaphor understanding and peer relationships) and, in particular, the effect of peer relationships on metaphor understanding, we found that being highly rejected, but not being less accepted, lowered later metaphor understanding. The effect of peer rejection on the development of metaphor

understanding is not surprising. Indeed, previous evidence reported that peer rejection was associated with poorer academic performances and cognitive development (Banerjee et al., 2011; Coie, 1990; Fite et al., 2013; Hinshaw, 1992; Masten et al., 2005). In addition, the negative effect of peer rejection on the later ability to understand a metaphor fits with the regulatory depletion model according to which individuals' cognitive inner resources are limited and are depleted by the regulation of negative emotions coming from stressful life events (Muraven & Baumeister, 2000; Muraven, Tice, & Baumeister, 1998). Since peer rejection is a stressful life event, it would limit children's development of cognitive functions, such as metaphor comprehension (Lecce et al., 2020). The present findings, in light of the regulatory depletion model, empathize the crucial role of stressful social experience (i.e., peer rejection), rather than the one of social exchanges (i.e., peer acceptance), in the development of cognitive functions. However, the lack of a significant predictive effect of peer acceptance on later metaphor understanding was unexpected in the light of socio-constructivist theoretical models (Clark, 2019; Piaget, 1932; Rubin et al., 2007; Tudge & Winterhoff, 1993; Vygotsky et al., 1978) and requires further investigations. Indeed, even if we found no effect of peer acceptance on the development of metaphor understanding, we do not think that peer acceptance has no effect at all on pragmatic skills. It is still possible that social exchanges (i.e., peer acceptance) would have a crucial role in the development of other pragmatic skills, such as metaphor production (see Clark, 2019).

Regarding the effect of metaphor understanding on social relationships, our results showed that having higher metaphorical skills lowered later probabilities of being rejected but did not enhance later probabilities of being accepted by peers. Notably, these results fit with empirical evidence on pragmatic conversational skills in pre-school children (see for a review, van der Wilt et al., 2019) and extend to middle childhood, and to a higher inferential pragmatic ability, namely metaphor understanding, the existence of a significant association between pragmatics and social relationships. In more detail, our results are in line with findings reported by van der Wilt and colleagues showing that children with poor pragmatic skills are less attractive playmates and, consequently, become highly rejected by peers (van der Wilt, van der Veen, van Kruistum, & van Oers, 2018; van der Wilt et al., 2020). Starting from this claim, we argue that even understanding the intended meaning during a conversation, especially in non-literal communicative exchanges (as in metaphors), is an ability that leads to being less attractive playmates, maybe because this skill is not only important but necessary for social relationships. In this view, not being able to understand metaphors would represent a risk factor for social relationships since it makes children less appealing social partners and consequently highly rejected. Strikingly, we found no significant predictive effect of metaphor understanding on later peer acceptance. This result is, apparently, in contrast with the literature claiming a social function of metaphors in adulthood (Bowes & Katz, 2015), but we think, this is not the case for two main reasons. First, while the literature on the social function in adulthood mainly adopted a dyadic approach to social

relationships focusing on adults interacting with one social partner, in this study we emphasized a group dimension. Interestingly, dyadic relationships differ from the group interactions in two ways: (1) dyadic relationships involve reciprocal feelings instead of unilateral perceptions about one individual from the entire group of peers (see for a review, Flannery & Smith, 2017); (2) the skills required for being accepted in a group are not necessarily the same for being a close and intimate partner in dyadic relationships (Flannery & Smith, 2017; Larson, Whitton, Hauser, & Allen, 2007). Second, in the present study, we examined the association between peer acceptance and the general ability to understand metaphors (measured using an offline task), while the literature on adulthood it was evaluated the prompting effect of metaphors on social relationships (presenting metaphors before the evaluation of closeness). For these reasons, we do not exclude a positive role of metaphor understanding, when presented before the interaction with a peer, on a dyadic interaction.

Crucially, we found that the longitudinal association between peer rejection and metaphor understanding had equal strength in the two directions, suggesting that the role of stressful social experiences in hindering the development of metaphor comprehension and the one of poor metaphorical skills in increasing the risk of being rejected by peers and are equally stronger.

Caveats and Conclusions

The present study is the first to investigate the longitudinal associations between metaphor understanding and peer relationships in middle-aged children.

Although interesting and pioneering the present study leaves some open questions for future studies. Indeed, future studies should explore the links between metaphor understanding and other measures of children's social relationships, such as friendship and dyadic likeability. This would help to investigate if dyadic relationships, differently from acceptance within the peer group, are related to metaphor understanding. This result is expected given the quasi-experimental study by Place and Backer (1991) that found a social partner to be more liked when displaying appropriate pragmatic skills during a dyadic conversation. One other gap in the present study is the lack of evaluation of social relationships by teachers. Indeed, on this matter, Cheung and Elliot recently found an association between pragmatics and children's likeability only when rated by teachers, suggesting that peer- and teacher- ratings convey different meanings of likeability, that are differently related to pragmatic skills (Cheung & Elliott, 2017). Similarly, it could be interesting to include a measure able to assess, in addition to metaphor understanding, the production of metaphors to investigate whether peer acceptance would promote the development of metaphorical skills at least in metaphor production, as it has been suggested by Clark (2019). Finally, future studies should also include other pragmatic measures to examine whether the associations with peer relationships in middle childhood generalized to high-level pragmatic skills besides metaphors.

Even if the present study leaves several open questions it also has several strengths. The first is that it is the first study to investigate the links between pragmatic and social relationships focusing on a high level inferential pragmatic

skill that is known to have a social role during adulthood (Bowes & Katz, 2015), namely metaphor understanding. Second, it uses the newly developed PMM task that was found to show good psychometric properties and can, therefore, be used in future studies. The third strength refers to the use of a measure of peer relationships that accounted for both peer acceptance and peer rejection, two indices that refer to different dimensions of social experience (Slaughter et al., 2015). The use of both indices allowed us to identify specific patterns of associations with metaphor understanding. Finally, the use of a longitudinal design allowed us to investigate the developmental associations between social relationships and pragmatics and to shed light on the bidirectional nature of the association between metaphor understanding and peer rejection.

Finally, given the bidirectional association between metaphor understanding and peer rejection, and given the literature showing a bidirectional relationship between peer rejection and ToM (Banerjee et al., 2011), the present work should prompt researchers to investigate the existence of the associations between metaphor understanding and ToM. This is the aim of the subsequent part of the present dissertation.

Chapter 2 – Theory of Mind and Metaphor Understanding in middle childhood

Abstract

The main aim of the present study was to investigate the association between individual differences in metaphor understanding and Theory of Mind (ToM) in typically developing middle-aged children. Preliminary, we were also interested in testing the psychometric properties of a short version of the Physical and Mental Metaphors (PMM) task, developed for assessing metaphor understanding. We assessed in 217 children (aged from 9 to 12 years) their vocabulary, working memory, socio-economic status, Theory of Mind, and metaphor understanding. In measuring metaphor understanding we decided to assess, not only metaphor comprehension but also metaphor interpretation, an index that allowed us to distinguish between physical and mental metaphors. Results showed that the short version of the PMM task is an adequate measure for capturing individual differences in metaphor interpretation, but not in metaphor comprehension, during middle childhood. In addition, we found that 9-year-olds (but not older children), who scored higher on ToM task, were also better in interpreting mental, but not physical, metaphors. We conclude that the link between metaphor and ToM is specific for the interpretation of mental metaphors and is stronger in early rather than later childhood.

Introduction

Theory of Mind (ToM) is the ability to attribute and make inferences about mental states in order to predict and explain others' behavior (Wimmer & Perner, 1983). Recent studies in this area of investigation expanded its developmental scope by showing significant changes in ToM well beyond pre-school years (Devine & Hughes, 2013; Dumontheil, Apperly, & Blakemore, 2010) and highlighted the existence of meaningful individual differences during middle childhood and pre-adolescence (Hughes & Devine, 2015). During school years, children become better at understanding the various aspects of complex social scenarios (Banerjee et al., 2011). In addition, during this developmental period, individual differences in ToM were found to be related to those in executive functions (Lecce et al., 2017), mental state conversations (Bianco, Lecce, & Banerjee, 2016), social relationships (Devine et al., 2016) and academic outcomes (Lecce, Caputi, & Hughes, 2011).

The extension to ToM research into middle childhood has allowed researchers to address unexplored issues, opening the way to the examination of new topics such as the relationships between ToM and advanced forms of social communication. The present study follows this recent trend in the literature by investigating the relationship between individual differences in ToM and in a specific aspect of pragmatics, namely metaphor understanding.

As already pointed out in Chapter 1, metaphor understanding reflects the ability to go beyond the literal meaning. This ability was investigated by some authors highlighting the difference between mental and physical metaphors

(Melogno et al., 2017; Vosniadou et al., 1984; Wang & Dowker, 2010; Winner et al., 1976). Surprisingly, the distinction between mental and physical metaphors has never been considered with respect to ToM. This is particularly striking given that mental and physical metaphors differently capitalize on the ability to attribute mental states, namely ToM. While the understanding of mental metaphors requires an inference about the mental states of the topic, the understanding of physical metaphors requires an inference about the physical aspects of metaphors' topic. Thus, distinguishing between metaphor types seems to be crucial to unpack the findings on the association between ToM and pragmatics in a broader sense (Matthews et al., 2018).

The role of ToM in metaphor comprehension was usually investigated in the literature on Autism Spectrum Disorder (ASD) showing conflicting results (Vulchanova, Saldaña, Chahboun, & Vulchanov, 2015; Whyte & Nelson, 2015). For example, several authors found impairments in metaphor comprehension in ASD (Kalandadze, Bambini, & Næss, 2019; Vulchanova et al., 2015), a clinical condition that is known to be characterized by a deficit in ToM (Baron-Cohen, 2000; Baron-Cohen, Leslie, & Frith, 1985). In line with this, some authors found that in this clinical condition, individuals who were better in ToM abilities also showed higher metaphor comprehension skills (Happé, 1993; Huang et al., 2015). These results drive to the conclusion that ToM is necessary to support metaphor comprehension (Happé, 1993). On the other hand, other researchers found that differences in metaphor understanding between ASD and typically developing children disappeared when language was taken under control, highlighting the

role of vocabulary abilities in metaphor comprehension (Kalandadze et al., 2016; Norbury, 2005). Thus, while earlier findings suggest an association between metaphor comprehension and ToM in ASD, the latter results posit that difficulties with metaphors are mainly due to verbal aspects rather than to ToM itself.

Given that the vast majority of studies focused on clinical samples, very little is known about typically developing children. Indeed, till now only one study looked at individual differences in ToM and metaphor understanding also in typically developing children (Whyte & Nelson, 2015). This study found that, in children aged 5 to 12 years, individual differences in nonliteral language, including metaphor comprehension skills, were related to those in ToM, even after controlling for language skills. However, this study was not focused on metaphor but considered the broader ability to understand nonliteral language. Nevertheless, this result, together with the findings reporting an association between metaphor understanding and peer rejection (see Chapter 1), and the ones showing a relationship between peer rejection and ToM (Banerjee et al., 2011), led us to expect a link between metaphor understanding and ToM.

The Present Study

The main aim of the present study was to assess the relationship between individual differences in ToM and metaphor understanding by adopting an individual differences approach and by distinguishing between physical vs. mental metaphors, though the PMM task, in different age groups. Preliminarily, we aimed (a) to test the psychometric properties of a short version of the Physical and

Mental Metaphors task and (b) to investigate similarities and differences in developmental changes of metaphor understanding and ToM in children aged 9 to 12 years. Regarding the preliminary aim (a), we decided to develop a shorter PMM task for two main reasons. On the one hand, it would be less cognitively challenging for children. On the other hand, a shorter scale, being easier to administer, would allow researchers to include this construct in large studies without having excessively long assessment sessions. Regarding the preliminary aim (b), we selected this age range since middle childhood is crucial for the development of both ToM and metaphoric understanding (see General Introduction).

Regarding the main aim, given that the accuracy score index did not allow to account for the differences between mental and physical metaphors (see Chapter 1), we decided to assess the distinction between mental and physical metaphors trough an interpretation score, that reflects how much children's answers refer to mental attributes. For example, a child who explains the mental metaphor "Soldiers are lions" by saying that "They are strong" shows a good comprehension of the metaphor, since she/he is able to retrieve the salient link between the topic and the vehicle, but scores low in mental interpretation since she/he does not refer to soldiers' mental states. Conversely, a child who answers "They are courageous" would score high both in comprehension and in mental interpretation, since she/he refers to soldiers' mental states. Interestingly, the distinction between mental and physical metaphors echoes the distinctions between inferences about the physical causes (control stories) and inferences

about mental states (ToM stories) tested in a classic ToM task such as the Strange Stories. Given that mental, but not physical, metaphors and ToM, but not control stories, require inferences about mental states, we expected a significant association between ToM (but not general inferential ability measured though physical stories of the Strange Stories task), and interpretation of mental (but not physical) metaphors. In addition, we expected this effect to be significant even after controlling for children's receptive language and working memory, which are known to be related to both metaphors (Carriedo et al., 2016; Norbury, 2005) and ToM (Hughes, 1998; Lecce et al., 2017). Finally, since cognitive skills become progressively specialized (Karmiloff-Smith, 1995), we expected the association between ToM and interpretation of mental metaphors to become weaker for older age groups.

Materials and Method

Participants

We recruited 217 participants aged 9 to 12 years through Northern Italy elementary and secondary schools. The sample was then split into four age groups: 62 9-year-olds (33 F, *M*age = 9 years; 6 months, *SD* = 0;3, age-range = 9;0 – 10;0), 48 10-year-olds (21 F, *M*age = 10;5, *SD* = 0;3, age-range = 10;0 – 11;0), 51 11-year-olds (26 F, *M*age = 11;5, *SD* = 0;3, age-range = 11;0 – 12;0), and 56 12-year-olds (20 F, *M*age = 12;4, *SD* = 0;3, age-range = 12;1 – 12;10). All children were fluent in Italian language and had no history of developmental delay or learning disorder. Assessing socio-economic status via the Family Affluence

Scale (FAS; Currie, et al., 2008), we found that 84.90% of the sample was classified as "high affluence" (range = 6–9), 14.70% as "middle affluence" (range = 3–5), and just one child (0.5%) as "low affluence" (range = 0–2).

Procedure

Before the children's assessment, we collected parental written consent for all participants. At the beginning of the school year, we assessed children's socio-economic status, verbal ability, working memory, metaphor understanding, and ToM. Children completed all the tasks individually in a quiet room at school during lecture time. This study has been approved by the local University Ethical Committee.

Measures

Socio-economic status. We measured socio-economic status via the Family Affluence Scale (Currie et al., 2008). This task required children to fill in a short questionnaire composed of 4 questions about the family wealth: (a) family car ownership (range = 0–2), (b) participants having/not having their own unshared room (range = 0–1), (c) number of computers at home (range = 0–3), and (d) number of times participants went on a holiday during the past year (range = 0–3). Total family affluence score was obtained by summing each response and it ranged from 0 to 9.

Verbal ability. Verbal ability was assessed via the Italian version of the vocabulary subtest of the Primary Mental Abilities (PMA; Rubini & Rossi, 1982; Thurstone & Thurstone, 1962). This task required children to find the synonyms of 30 target words, choosing among four alternatives, in 7 mins maximum. Total scores ranged from 0 to 30.

Working memory. We measured working memory through the Backward Digit Span task of the WISC-R Italian version (Orsini, 1993). During the task, children, after having listened to a series of digit sequences, had to recall them in the reverse order. Children completed seven sequences of digits of increasing difficulty. Indeed, the number of digits in each sequence increased from two to eight. A total working memory score was calculated by summing the number of correct sequences, and it ranged from 0 to 7.

Theory of Mind. Theory of Mind was evaluated via the Strange Stories task (Happé, 1994). The task is composed of stories depicting social situations and requires children to interpret the main character's utterances by making inferences about his/her behavior. We selected three physical stories, that require inferences about the physical causes of the character's behavior and five mental stories, that require inferences about the character's mental states (two double bluffs, two misunderstandings, one persuasion, and one white lie).

After reading each story, participants had to answer an open-ended question in which they had to explain the reasons underlying the main character's

behavior in a written form. No time limit was imposed. In line with scoring guidelines (White, Hill, Happe, & Frith, 2009), we rated children's answers on a 3-point rating scale: 0 for incorrect and "Don't know" answers, 1 for partially correct and implicit answers, and 2 for full and explicit answers. Total scores were calculated by summing the score obtained at each story and could range from 0 to 6 for physical stories and from 0 to 10 points for mental stories. A second rater coded independently 25% of the responses and we found an almost perfect inter-rater agreement assessed through Cohen's kappa (k = .85).

Metaphor Understanding. Metaphor understanding was evaluated through a short version of the Physical and Mental Metaphors task (see Chapter 1). This task is composed of three physical metaphors and three mental metaphors casually selected from the two sets described in Chapter 1. All nominal metaphors are non-lexicalized and follow the prototypical structure "X is Y" in which the topic (X) is always a person while the vehicle (Y) is always a non-human entity. The physical metaphors were: (1) "Le ballerine sono farfalle" (literal translation "Dancers are butterflies"), (2) "Gli scalatori sono scoiattoli" ("Climbers are squirrels"), and (3) "I giocatori sono elefanti" ("Players are elephants"). The mental metaphors were: (4) "I Soldati sono leoni" (literal translation "Soldiers are lions"), (5) "Il papa è un vulcano" ("Daddy is a volcano"), and (6) "la maestra è un ghiacciolo" ("The teacher is an icicle"). Physical metaphors require inferences about the physical/behavioral features of the topic, whereas mental metaphors require inferences about the mental states of the topic.

To verify that even these shorter sets were balanced for lexical frequency and familiarity, differing only in the preferred interpretation (either physical or mental) we repeated the analyses performed in Study 1 and described in Chapter 1. For a full description PMM items see Table 1.1s and for a full description of the ratings and for the comparison between the two sets of the PMM short version see Table 2.1. Results were the following: (1) there were no differences in familiarity, assessed via a frequency dictionary of child language (Marconi et al., 1994), between the target words in the physical set and the ones in the mental set (Table 2.1); (2) familiarity and aptness, checked with a 7-point Likert-type rating scale administered to 53 young adults, did not differ across sets (Table 2.1); (3) mental and physical metaphors were different in the preferred interpretation when assessed both via a categorical task and a parametric task (see Measures section in Chapter 1 for a description of the two tasks). Results from the categorical task showed that for metaphors in the physical set, the agreement on physical interpretation was > 95%, and for metaphors in the mental set the agreement on mental interpretation was > 90%, with highly significant chi-squared for each set (Table 2.1) and each metaphor, $X^2(1)_s > 21.13$ $p_s < .001$.

Table 2.1. Comparison between the mental set and the physical set of the extended version of the Physical and Mental Metaphors task.

		All materials	Physical Set	Mental Set	Statistics
Frequency		74.38 (54.77)	85.83 (10.42)	62.92 (83.64)	$t(2.06) = 0.47$, $p = .683$
Familiarity		3.84 (0.85)	3.65 (1.25)	4.04 (0.35)	$t(4) = -0.53$, $p = .625$
Aptness		4.43 (0.91)	4.30 (1.36)	4.55 (0.43)	t(4) = -0.30, p = .781
Categorical task	Agreement on physical interpretation	–	95.83% [90-100]	4.17% [0-9]	$\chi^2(1) = 88.33$, p < .001
	Agreement on mental interpretation	–	8.33% [6-9]	91.67% [90-93]	$\chi^2(1) = 81.33$, p < .001
Parametric task	Physical attributes	4.51 (1.23)	5.46 (0.85)	3.56 (0.63)	$t(4) = 3.12$, $p = .036$
	Mental attributes	3.32 (1.85)	1.64 (0.25)	5.00 (0.27)	$t(4) = -15.87$, $p < .001$
	Difference score	–	3.82 (1.07)	1.43 (0.67)	$t(4) = 3.29$, $p = .030$

Notes. For frequency, familiarity, aptness and the parametric task scores, the Table reports mean and SD in parenthesis. These ratings were obtained from 53 subjects (40 F; *M*age = 23.91, *SD* = 2.33, age-range = 21 – 32; *M*education = 15.83, *SD* = 1.67, education range = 13 – 18). For the categorical task scores, the Table reports mean agreement and range in parenthesis. These ratings were obtained from 32 subjects (22 F; *M*age = 22.35, *SD* = 1.99, age-range = 21 – 30; *M*education = 15.16, *SD* = 0.95, education range = 14 – 20).

Results from the parametric task showed that metaphors in the physical set scored significantly higher on physical attributes compared to metaphors in the mental set and that metaphors in the mental set scored significantly higher on mental attributes compared to metaphors in the physical set. Interestingly, as expected, the difference between the mean score on the physical attributes and the mean score on the mental attributes, was significantly bigger for metaphors in the physical set compared with the mental set. Overall, these data confirm the findings of the previous study (see Chapter 1) even for the short version of the PMM task, supporting the distinction between the physical and the mental sets and indicating that, differently from metaphor in the physical set, metaphors in the mental set were also open to physical interpretations.

After an example item that was considered together with the experimenter, children were required to explain orally the meaning conveyed by each metaphor. Children's answers were coded, following the scoring guidelines (see Chapter 1), according to their level of accuracy (defined as the ability to articulate the link between the topic and the vehicle). 0 points were assigned for "Don't know", literal, or incorrect answers. 1 point was assigned for incomplete answers or answers referring to a non-salient feature of the metaphor vehicle. 2 points were assigned for complete answers that refer to the salient features of the metaphor vehicle. In addition to the level of accuracy, children's answers were coded according to their level of interpretation (physical vs. mental). 0 points were given to answers referred to physical attributes (for the metaphor "I Soldati sono leoni" [Soldiers are lions] "Sono forti" [They are strong]) or actions ("Corrono veloci"

[They run fast]) of the topic. 1 point was given to answers referred to psychological attributes of the topic ("Sono coraggiosi" [They are courageous]).

A second rater independently coded 25% of the children's answers and interrater agreement was established via Cohen's kappa, indicating an almost perfect for each index (k = .96 and k = .88 for the physical set accuracy and interpretation, respectively. k = .98 and k = .94 for the mental set accuracy and interpretation, respectively).

The new distinction between the level of accuracy and interpretation allowed us to take fine-grain analyses on interpretation rather than simply on the ability to understand a metaphor. Specifically, the interpretation score allowed to consider cases in which children gave plausible, yet physical, interpretations of metaphors in the mental set, and vice versa.

Statistical analyses plan

Before testing our main hypothesis, that is, the existence of an association between ToM and metaphor understanding, we conducted preliminary analyses and evaluated the developmental changes in all study variables. Regarding preliminary analyses, we first checked the level of accuracy in children's answers for all items and the existence of different interpretations across the physical and mental sets. Regarding the developmental changes, we performed a series of ANOVAs on the control (verbal ability and working memory) and the focus variables (ToM and metaphor understanding), with age groups (four levels) as the between-subject variable.

Regarding our main aim, we performed a series of correlation and partial correlation analyses (controlling for verbal ability, working memory, and socio-economic status) between ToM and metaphor understanding.

Results

Preliminary analysis

A preliminary inspection at children's metaphor accuracy revealed that the great majority of children scored high on this index, with 98.2% and 92.2% of children showing a performance that was equal or bigger than 3 on both the mental and the physical metaphor set. For this reason, we decided to focus, in the main analysis, on metaphor interpretation rather than accuracy. Preliminarily, to check the soundness of this task in distinguishing between mental and physical metaphors when focusing on interpretation, we compared the interpretation score between mental and physical metaphors. We found that children scored higher in interpretations of mental metaphors compared to physical ones, $t(216) = 32.88$, $p = .000$. Thus, as expected, children were more likely to refer to mental states when interpreting a mental rather than a physical metaphor. In addition, we found no significant correlation between the interpretation of physical and mental metaphors, $r = -.04$, $p = .56$, indicating that inference about mental states was specific for those metaphors requiring it.

Developmental effects

Descriptives for each age group are reported in Table 2.2.

Table 2.2. Descriptive Statistics for each age group.

	Mean (SD)			
	Age 9	Age 10	Age 11	Age 12
Control Variables				
Verbal Ability	15.56 (5.49)	21.19 (6.39)	22.67 (7.44)	25.32 (6.64)
Working Memory	2.52 (0.83)	2.92 (0.99)	3.12 (0.93)	3.11 (0.80)
Metaphors Task				
Physical Set – Accuracy	4.45 (1.36)	4.60 (1.18)	4.24 (1.41)	4.59 (1.33)
Physical Set – Interpretation	0.03 (0.18)	0.08 (0.28)	0.14 (0.35)	0.5 (0.23)
Mental Set – Accuracy	4.98 (1.17)	5.31 (1.03)	5.16 (0.92)	5.25 (1.0)
Mental Set – Interpretation	1.68 (0.76)	1.90 (0.75)	2.08 (0.72)	2.04 (0.84)
ToM Task				
Mental Strange Stories	6.39 (1.37)	7.15 (1.66)	7.53 (1.47)	7.44 (1.48)
Physical Strange Stories	3.06 (1.30)	3.19 (1.50)	3.63 (1.67)	3.66 (1.46)

We assessed the age differences in all the study variable through a series of ANOVAs. Regarding vocabulary we found a main effect of age, $F(3,213) = 23.98$, $p < 0.001$, $\eta_p^2 = .25$. In detail, post-hoc Bonferroni comparisons revealed that 9-year-olds scored lower in verbal ability compared to the other age groups and that 10-year-olds scored lower in verbal ability compared to 12-year-olds. No other contrast reached the statistical significance.

Regarding working memory, again we found a significant age effect, $F(3,213) = 5.98$, $p = .001$ $\eta_p^2 = .08$ with 9-year-olds scoring significantly lower compared to 11- and 12-year-olds. No other contrast reached the statistical significance.

Regarding ToM, we found a main effect of age on mental, $F(3,213) = 7.09$, $p < .001$, $\eta_p^2 = .09$, but not on control stories of the Strange Stories task, $F(3,213) = 2.36$, $p = .07$, $\eta_p^2 = .03$. In detail, Bonferroni comparisons showed that 9-year-olds scored lower in ToM compared to other age groups. No other contrast reached the statistical significance.

Regarding metaphor interpretation, our analyses showed a main effect on mental, $F(3,213) = 3.59$, $p = .01$, $\eta_p^2 = .04$, but not physical metaphors, $F(3,213) = 1.67$, $p = .18$, $\eta_p^2 = .02$. Post-hoc Bonferroni comparisons showed that 9-year-olds scored lower in mental interpretation compared to 11- and 12-year-olds. No other contrast reached the statistical significance.

Associations between ToM and metaphors

Pearson's correlations between individual differences in interpreting mental metaphors, physical metaphors, and the Strange Stories task are reported in Table 2.3 for each age group.

We found that individual differences in interpreting mental, but not physical, metaphors were significantly correlated with those in ToM (assessed via the mental stories of the Strange Stories task) in 9-year-olds, but not in any other age group. Crucially this association persisted even after controlling for verbal

ability, working memory, and socio-economic status (Table 2.3). Interestingly, this association remained significant also when we controlled for accuracy in the understanding of mental metaphors, in addition to the control variables cited above, $r = 0.32$, $p = .01$. Even if we found a significant association between the interpretation of mental metaphors and control stories of the Strange Stories task in 9- and 11-year-olds, this correlation fell below the significance level when we controlled for verbal ability, working memory, and socio-economic status (Table 2.3).

Overall, we found only in 9-year-olds a specific association between metaphor interpretation and strange stories for mental items. Indeed, we found no significant correlation between the interpretation of mental metaphors and physical control stories of the Strange Stories task and, on the same lines, no significant association between the interpretation of physical metaphors and mental stories of the Strange Stories task.

Table 2.3. Correlations and partial correlations (in parenthesis) between ToM and Metaphor Task for each age group.

		Interpretation of physical metaphors	Physical Strange Stories	Mental Strange Stories
Interpretation of mental metaphors	Age 9	.08 (.04)	.29* (.26)	.37** (.34*)
	Age 10	-.05 (.03)	.25 (.10)	.27 (.19)
	Age 11	-.21 (-.13)	.29* (.23)	.24 (.14)
	Age 12	.02 (.04)	.18 (.08)	.13 (.09)
Interpretation of physical metaphors	Age 9		.06 (.02)	.02 (-.10)
	Age 10		-.09 (-.14)	-.03 (-.03)
	Age 11		-.04 (.01)	-.15 (-.07)
	Age 12		.01 (.03)	.21 (.27)
Physical Strange Stories	Age 9			.25* (.20)
	Age 10			.39** (.22)
	Age 11			.31* (.31*)
	Age 12			.26* (.09)

Notes. + $p < .06$. * $p < .05$. ** $p < .01$. *** $p < .001$. Partial correlations controlling for verbal ability, working memory, and socio-economic status are given in parenthesis.

Discussion

This study investigated the associations between ToM and metaphor understanding in different age groups by distinguishing between mental and physical metaphors. Before discussing the main results, we would like to highlight our preliminary findings concerning the psychometric properties of the short version of the PMM task and the developmental changes in metaphor comprehension and ToM across middle childhood. Our findings showed that all children scored high in accuracy and that 9-year-olds performed lower compared to other age groups (who did not differ significantly one from another) in the interpretation of mental, but not physical metaphors. These results suggest that the short version of the PMM task is an adequate measure for detecting individual differences in metaphor interpretation but not in metaphor comprehension. Focusing on the development of ToM, we found that 9-year-olds performed significantly lower compared to other age groups (who did not differ significantly one from another) in ToM. Taking together the developmental findings, we observed a parallel in developmental timing, with the age of 9 years as a key turning point for the development of both ToM and interpretation of mental metaphors. This mirroring between ToM and metaphor interpretation development suggests that there may be a relationship between ToM and interpretation of psychological attributes of metaphors.

This consideration brings us to the main aim of the present study (i.e., the existence of an association between ToM and metaphor interpretation in middle-aged typically developing children). We found that in 9-year-olds (but not in older

children) ToM was related to the interpretation of mental, but not physical metaphors, independently of children's verbal ability, working memory, and socio-economic status. The fact that 9-year-olds who are better in ToM are also more likely to better interpret mental, but not physical, metaphors, suggests that:

(1) The association between metaphor interpretation and ToM is not general, but rather specific for those metaphors whose interpretation requires an inference about mental states;

(2) Mental metaphor interpretation is specifically associated with the ability to make inferences about mental states (i.e., ToM assessed via the mental stories of the Strange Stories task). This association does not extend to the ability to make inferences about physical states (assessed via the physical stories of the Strange Stories task);

(3) The specific association between the interpretation of mental metaphors and ToM changes across development, being stronger in earlier developmental phases.

These three findings have important implications.

First, the specific association between ToM and interpretation of mental, but not physical, metaphors support the theoretical distinction between psychological and physical metaphors. Thus, the kind of inference to be drawn (i.e., physical vs. mental) might indeed play a role in modulating the involvement of the ToM system. Our data suggest that, even if pragmatic interpretation always requires a certain degree of attribution of intentions (Sperber & Wilson, 2002), beyond this basic level of the communication process, ToM involvement might

vary depending on metaphor type. In detail, focusing on metaphors, when the relevant properties of the topic refer to mental aspects (e.g., "Daddy is a volcano"), the ToM load is likely to increase in typically developing children. A similar view is supported by a recent study on ASD children, where cognitive training was found to be more effecting in improving the comprehension of sensory, rather than psychological, metaphors (Melogno et al., 2017), suggesting that the two types of metaphors are explained, at least in part, by different mechanisms.

Second, the stronger relationship between metaphor and ToM in earlier ages offers a novel developmental angle to the debate over the relation between metaphor (and even the broader issue of pragmatics) and ToM. This result fits with the literature about the relationship between ToM and other non-literal language aspects, since it is usually based on data coming from young children (Caillies & Le Sourn-Bissaoui, 2008). Further evidence, in line with the present result, comes from the literature on the association between ToM and language (receptive and expressive). Indeed, there are studies focusing on the link between ToM and the language, showing that this association is stronger in pre-schoolers (Milligan, Astington, & Dack, 2007) and becomes weaker across middle childhood (Lecce, Zocchi, Pagnin, Palladino, & Taumoepeau, 2010). From a more general point of view, the present result is in line with Karmiloff-Smits's model of progressive modularization and specialization of cognitive skills (Karmiloff-Smith, 1995).

Caveats and Conclusions

The present study is the first that explores the relationship between ToM and metaphor understanding, distinguishing between mental and physical metaphors. However, despite the novelty of the present finding, our study is not free from limitations.

First, the short version of the PMM task was found not to be able to detect individual differences, especially when focusing on accuracy. Future studies should employ a larger number of stimuli than the ones used in the present study and should elucidate the role of ToM in metaphor understanding even when accuracy, in addition to interpretation, is considered.

Second, correlational analyses can only suggest an association, but cannot support claims on developmental relationships and causality. Future studies should investigate also the nature of the relationship between ToM and the understanding of mental vs. physical metaphors, possibly with both longitudinal and training studies. Indeed, while longitudinal studies demonstrate genuine developmental relationships, training studies verify if these genuine relations have also a causal nature (Bradley & Bryant, 1983). These two study designs would indeed enable us to understand if ToM predicts and promotes the development of pragmatics in typically developing children or if, vice versa, pragmatics predicts and promote the development of ToM.

Despite these limitations, our study presents novel findings that contribute to clarify the relationship between ToM and pragmatics in middle childhood, with two main strengths. The first is the distinction between two types of metaphors,

physical and mental, that require inferences about physical and mental states, respectively. The second is the developmental focus of the study, which covers a wide age range and showed that: (a) the development of mental metaphor interpretation mirrors the development of ToM, and (b) the association between ToM and the interpretation of mental metaphors is stronger in earlier developmental stages.

Chapter 3 – Theory of Mind and Metaphor Understanding: longitudinal associations in middle childhood

Abstract

The main goal of the present research was to investigate the developmental relationships between Theory of Mind (ToM) and metaphor understanding in middle-aged typically developing children using a short-term longitudinal design. In addressing this topic, we adopted a fine-grained analysis and distinguished, between (a) metaphor accuracy (the ability to find a link between metaphor topic and vehicle) and specificity of mental interpretation (the ability to interpret mentally mental but not physical metaphors) using the 12-items version of the Physical and Mental Metaphors (PMM) task, and (b) ToM and the ability to make inferences about physical states. We tested 54 typically developing children at baseline (age 8;6 - 9;4 years) and 6 months later for vocabulary, inferential skills about mental (ToM) and physical states, and metaphor understanding (accuracy and interpretation). Preliminary results on the soundness of the extended version of the PMM task showed that it is an adequate measure for capturing individual differences in metaphor comprehension and specificity of mental interpretation, during middle childhood. In addition, we found that inferential skills about physical states were bi-directionally linked to the ability to comprehend a metaphor and that specificity of mental interpretation predicted later ToM (and not the reverse). We conclude that metaphor comprehension and general inferential ability support each other during development and that the tendency to mentally interpret mental, but not physical, metaphors drive ToM development.

Introduction

As we already posited in the previous chapter, the associations between ToM and metaphor understanding are largely debated in the literature. On the one hand, on a theoretical basis, the associations between these two skills are expected, since metaphor understanding requires the ability to make an inference about the speaker's intentional state (Sperber & Wilson, 1995). On the other hand, empirical studies on atypical populations showed conflicting results (see Introduction section in Chapter 2).

The literature on typically developing children about the relationships between children's ability to understand the meaning conveyed by a metaphor and the ability to infer the mental states underlying others' social behavior, even if scarcer, shows higher agreement. Indeed, evidence from a cross-sectional study in middle childhood reported a relationship between individual differences in nonliteral language, including metaphor comprehension skills, and those in ToM, even after controlling for basic language abilities (Whyte & Nelson, 2015). In addition, results from the study presented in Chapter 2 showed that 9-year-olds' individual differences in the interpretation of mental metaphors significantly correlated with those in the ability to make inferences about mental (i.e., ToM) but not physical states (see Chapter 2). Notably, this association was significant even over and above individual differences in verbal abilities, working memory, and socio-economic status.

Despite this evidence, no study has yet investigated the longitudinal associations between ToM and metaphor understanding. Reasons to expect a developmental link between these two skills during middle childhood in typically developing children came from literature reporting parallelism in the developmental timing of these two skills. The improvement in the ability to understand metaphors from pre-school over primary school years (Noveck et al., 2001; Rundblad & Annaz, 2010; Vosniadou et al., 1984; Winner et al., 1976) mirrors the increase in children's ability to infer the correct mental states in several social situations (Lagattuta, Elrod, & Kramer, 2016; Lecce et al., 2017).

However, based on existing experimental evidence and given the cross-sectional nature of studies exploring the association between ToM and metaphor understanding, we are not allowed to drive any conclusion about the directionality of the link between ToM and metaphor understanding.

The Present Study

The present study adopted a longitudinal cross-lagged designed to examine not simply the existence, but also the direction of the associations between individual differences in ToM and the ones in metaphor understanding. Therefore, we want to fill the gap in the literature, shedding light on the debate about the longitudinal relationships between ToM and metaphor skills and helping to tease apart different theoretical hypotheses on the causal role of ToM on pragmatics and vice versa. Capitalizing on the study described in the previous

chapter (see Chapter 2) we decided to maintain the same tasks for measuring ToM and metaphor understanding.

Therefore, regarding the ToM task, we adopted again the Strange Stories task, since it is made up of mental and physical stories and allows to distinguish between children's ability to make inferences about mental states and the more general ability to make inferences from implicit physical information. Regarding the metaphorical task, we employed again the Physical and Mental Metaphors (PMM) task, since it distinguishes between two types of metaphors, mental and physical, which differently capitalize on the ability to attribute mental states. Crucially, the PMM task allows tackling different aspects of metaphor understanding, considering not only the more general ability to comprehend the message conveyed by the metaphor but also the ability to interpret a metaphor as referring to mental states. This double level of analysis is reflected in two different scores: accuracy and interpretation, respectively. Interestingly, in the study described in Chapter 2, we found that ToM was associated with the interpretation of metaphors involving mental, but not physical, aspects. This finding suggests that ToM skills are particularly triggered by some, and no other, metaphors, prompting future research to maintain this distinction when focusing on interpretation. However, in the study of Chapter 2, we were not able to perform any analysis on metaphor accuracy due to a ceiling effect. Thus, we decided to use the extended version of the PMM task (instead of the short version used in Study 2), in order to have an instrument that would be more sensitive to individual differences in middle childhood. For the same reason, we also decided to increase

the possible range of scores of the interpretation coding scale (see Materials and Method). In addition, we computed a more accurate index of children's ability to interpret a metaphor, called specificity of mental interpretation, which allowed us to control for children's errors. This index reflects children's ability to interpret a metaphor as conveying mental features of the topic only when appropriate, that is when the metaphor is indeed referring to mental and not to physical features (see Materials and Method).

The aims of the current study can be summarized in four points.

The first aim was to examine, for the first time, the development of ToM and metaphor understanding (both accuracy and specificity of mental interpretation) over a 6-months period in middle childhood.

The second aim was to verify the longitudinal rank-order stability of ToM and metaphor understanding over 6 months in middle childhood.

The third aim was to examine the concurrent relationships between ToM and metaphor understanding in middle childhood.

The fourth and main aim was to assess the developmental associations between metaphor understanding and ToM skills.

Concerning these aims, we had the following hypotheses.

First, based on the studies that found a developmental progression during middle childhood in both ToM and metaphor understanding (Austin, Groppe, & Elsner, 2014; Lecce et al., 2017; Pinto, Melogno, & Iliceto, 2011; Winner et al., 1976), we expected to find an improvement in ToM and metaphor understanding from Time 1 to Time 2.

Second, we expected to find high rank-order stability in the performance on ToM and metaphors tasks.

Third, we hypothesized to find concurrent relationships between ToM and metaphor understanding, with different patterns for the two scores of the PMM, that are, the one indexing the ability to comprehend a metaphor (accuracy) and the one indexing the tendency to give a mental interpretation to a mental, but not a physical metaphor (specificity of mental interpretation). In line with the results of Study 2 (see Chapter 2), we expected to find a specific concurrent relationship between ToM (but not the ability to make inferences about physical states) and specificity of mental interpretation of metaphors. Conversely, we expected to find an association between ToM and metaphor accuracy that was not modulated by the metaphor type. Our expectations were based on previous evidence of an association between children's ToM and general metaphorical ability, regardless of the distinction between mental and physical metaphors (Happé, 1993; Huang et al., 2015; Norbury, 2005), and on studies that found that mental and physical metaphors are equally difficult to understand (Nippold et al., 1984).

Fourth, regarding the main aim of the present study (i.e., to investigate the direction of the relationships between ToM and metaphor understanding), existing literature did not allow us to formulate a clear hypothesis, making the present study explorative in nature.

Starting from Study 2, that found a significant cross-sectional relationship between ToM and metaphor understanding in middle childhood (see Chapter 2), three patterns of results can be expected: (1) ToM has a role in the development of

metaphorical skills (both accuracy and specificity of mental interpretation) and, therefore, ToM at Time 1 predicts metaphor accuracy and specificity of mental interpretation at Time 2; (2) metaphor understanding (both accuracy and specificity of mental interpretation) has a role in ToM development and, therefore, metaphor accuracy and specificity of mental interpretation at Time 1 predict ToM at Time 2; (3) the direction of the developmental associations between ToM and metaphor understanding varies depending on the pragmatic aspect considered, namely accuracy vs. specificity of mental interpretation.

The first hypothesis (e.g., ToM predicts Metaphor understanding) fits with studies conducted in the theoretical framework of the Relevance Theory, according to which pragmatics is a sub-module of ToM evolved and dedicated for human communication (Happé, 1993; Sperber & Wilson, 2002). This hypothesis is also in line with several position papers arguing that early forms of mentalizing, such as pretense and perspective-taking, are precursors of figurative language use, including metaphors (Clark, 2019; Falkum, 2019).

The second hypothesis (e.g., metaphor understanding predicts ToM) fits with the broader literature dealing with the relationship between language abilities and ToM. Indeed, longitudinal studies on middle childhood reported that early language abilities predicted later ToM, measured via the Strange Stories task (Devine et al., 2016; Lecce et al., 2017). Accordingly, Zufferey stated that while in the first year of life ToM enables lexical acquisition, later on in development, language acquisition provides children a tool for reasoning about others' mental states and for developing advanced forms of ToM (Zufferey, 2010).

Finally, the third hypothesis (e.g., the direction of the developmental relationships between ToM and metaphor differs for accuracy and specificity of mental interpretation) fits with literature stating that the relationships between pragmatics and ToM vary depending on several different factors. For instance, Andrés-Roqueta & Katsos (2017) claimed that pragmatic skills are differently engaged depending on the specific communicative situation. In detail, they distinguished between linguistic-pragmatics and social-pragmatics. While the former is used in situations that require structural language rather than ToM competence, the latter is used in circumstances that largely capitalize on ToM skills. A similar perspective was adopted by Matthews and colleagues (2018) that in a recent review stated that the link between ToM and pragmatics is not always found because (a) ToM may be necessary only for some and not for all pragmatic skills, and (b) their relationships are better explained by more general cognitive skills (e.g., working memory and language). In line with these claims and this third hypothesis, we expected to find that ToM is more likely to be involved in specificity of mental interpretation rather than in metaphor comprehension. Indeed, metaphor comprehension is likely to be related to a broader inferential ability rather than to ToM per se. Thus, we expected to find a developmental relationship (a) between specificity of mental interpretation and ToM, and (b) between metaphor accuracy and inferential skills as measured by the control stories of the Strange Stories task. In line with this hypothesis, both metaphor understanding and ToM were reported to require some inferential skills. On one hand, during metaphor understanding the speaker's intended meaning is grasped

via an inference from the encoded linguistic meaning (e.g., Sperber & Wilson, 1995, 2005) and metaphorical tasks, such as the PMM task, require children to make pragmatic inferences based of the phrasal context. On the other hand, ToM is defined as the ability to infer others' mental states, and ToM tasks, including the Strange Stories task, require children to make pragmatic inferences about characters' mental states. In line with this view, both metaphor comprehension and ToM in middle childhood are related to tasks of reading comprehension that include inferential questions (Cantin, Gnaedinger, Gallaway, Hesson-McInnis, & Hund, 2016; Seigneuric, Megherbi, Bueno, Lebahar, & Bianco, 2016).

To explore these hypotheses, we designed a two-wave cross-lagged longitudinal study in which we measured individual differences in children's ToM and metaphor understanding across two time-points from 8.86 to 9.36 years of age. We selected this age range following previous results showing that ToM and metaphor interpretation are linked in 9-, but not in 10-year-old children (see Chapter 2). In addressing these issues, we also controlled for rank-order stability, for the general ability to make inferences (measured via the physical control stories of the Strange Stories task) and for verbal abilities, which are known to be related to both ToM and pragmatics (Matthews et al., 2018; Milligan et al., 2007).

Materials and Method

Participants

We recruited 54 typically developing school-aged children (31 F, *M*age = 8 years; 10 months, *SD* = 0;3, age-range = 8;6 – 9;4, at Time 1) from 4 classes of

two public primary schools located in Northern Italy. All children spoke Italian at home and have no history of specific learning disorders or developmental delays.

Procedure

At the beginning of the study, parental written consent was collected for all children. We assessed children's vocabulary, ToM, and metaphor understanding at baseline (Time 1, May) and 6 months later (Time 2, November). At each time point, vocabulary was collected in a whole-class session and ToM and metaphor understanding were assessed in an individual session taking place in a quiet room located in the school. This study has been approved by the local University Ethical Committee.

Measures

Vocabulary. We measured vocabulary through the vocabulary subtest of the Primary Mental Abilities (PMA; Rubini & Rossi, 1982; Thurstone & Thurstone, 1962). Children were asked to find, in seven minutes maximum, the synonyms of 30 target words choosing among four alternatives. The total score was given by the number of correct answers (range = 0–30).

Theory of Mind. Theory of Mind was measured via the Strange Stories task (Happé, 1994). Children were required to interpret the main character's utterances by making inferences about the character's behavior. We administered seven mental stories (two double bluffs, two misunderstandings, two persuasions,

and one white lie) and six physical/control stories. Mental stories require children to make inferences about mental states, whereas physical stories require children to make context-appropriate inferences about the mechanical and the physical cause of the character's behavior.

After reading each story, the experimenter asked children to answer an open-ended question in which they had to explain the reasons underlying the main character's behavior in a written form. Following the scoring guidelines (White et al., 2009), 0 points were assigned for incorrect and "Don't know" answers, 1 point for partially correct and implicit answers, and 2 points for full and explicit answers. Total scores were given by the sum of the scores obtained at each mental (range = 0–14) and physical (range = 0–12) story of the Strange Stories task. A second rater independently coded 25% of the responses at each time point, and inter-rater agreement was established using Cohen's kappa, reporting an almost perfect agreement at both time points (at Time 1, k = .84 and k = .90 and at Time 2, k = .95 and k = .91 for mental and physical stories of the Strange Stories task, respectively).

Metaphor Understanding. We evaluated metaphor understanding via the extended version of the Physical and Mental Metaphors (PMM) (see Chapter 1). This task is composed of six mental and six physical nominal metaphors in the form "X is Y". To be understood, physical metaphors require inferences about the physical/behavioral features of the topic, while mental metaphors require inferences about the mental states of the topic. The two sets were equal in terms of

lexical frequency of the target words and of metaphors familiarity (see Chapter 1). During the task, the experimenter asked children to explain orally the meaning conveyed by each metaphor, after an example item considered together.

In the present study, we used the extended version (see Chapter 1) instead of the short version of the PMM task (see Chapter 2) to increase score variability and, thus, to allow for a better investigation of individual differences even for the accuracy score (see Results section in Chapter 2). For a full description of the ratings and items of the PMM task see Table 1.1a and Table 1.1b.

Children's answers were coded according to the level of (a) accuracy, defined as the ability to articulate the link between the topic and the vehicle, and (b) interpretation, defined as the ability to interpret metaphors as mental or physical.

Following the scoring guideline (see Chapter 1), accuracy was coded on a 3-point scale: 0 was assigned for incorrect, literal, and "Don't know" answers. A score of 1 was assigned for answers that were incomplete or referred to non-salient features of the metaphor topic, and 2 for answers that were complete and referred to salient features of the metaphor topic. We calculated inter-rater agreement through Cohen's kappa on the 25% of the answers, finding an almost perfect agreement at both time points (k = .89 and k = .83 for Time 1 and Time 2, respectively). A total accuracy score was calculated by summing the scores obtained at each metaphor (range = 0–24).

Interpretation scores were attributed on a 4-point scale, and not on a 2-point scale as in Chapter 2, in order to increase scores variability and to take into

account the ambiguity given by the fact that some words could vehicle either mental or physical meaning. A score of 0 was given for "Don't know" answers, 1 for answers referred to physical attributes (e.g., for the metaphor "La nonna è una colonna" [Grandma is a column], "È molto alta" [She is very tall]) or actions ("Sta ferma" [She stands still]) of the metaphor topic, 2 for ambiguous answers that could be referred either to psychological or physical attributes of the metaphor topic ("È importante" [She is important]), and 3 for answers referred to psychological attributes of the metaphor topic ("È una persona su cui mi appoggio" [She is someone I can count on]). We calculated inter-rater agreement through Cohen's kappa on the 25% of the answers, finding an almost perfect agreement at both time points (k = .88 and k = .87 for Time 1 and Time 2, respectively). We computed an index of children's ability to interpret a metaphor, namely specificity of mental interpretation, by subtracting the interpretation score of physical metaphors from the interpretation score of mental metaphors (range = −18−18). This index reflects children's ability to interpret mentally only mental and not physical metaphors and, thus, is a more accurate and complete index than the two separate scores (interpretation of mental metaphors and interpretation of physical metaphors). Indeed, specificity of mental interpretation controls for hyper-mentalizing errors when a metaphor that refers to physical aspects is interpreted mentally.

Statistical analyses plan

Before investigating the main aim of the present study (i.e., the direction of the associations between ToM and metaphor understanding in middle childhood) we conducted a series of preliminary analyses:

(1) We examined, via a series of paired sample t-tests, the developmental changes in metaphor accuracy and specificity of mental interpretation. In addition, we ran a series of repeated-measures ANCOVAs controlling for vocabulary to investigate whether the developmental changes persist even after controlling for vocabulary;

(2) We checked for the stability of the key variables performing correlations and partial correlations (controlling for vocabulary at Time 1);

(3) We examined the concurrent relationships between ToM and metaphor understanding (both accuracy and specificity of mental interpretation) performing within-time correlations and partial correlations (controlling for concurrent vocabulary).

In the main analyses, we investigated the directionality of the longitudinal links between ToM and metaphor understanding by performing across-time correlations, partial correlations (controlling for concurrent vocabulary), and hierarchical regressions. In the hierarchical regression analyses, we forced the entry of control variables (vocabulary at Time 1 and the stability of the dependent variable) at Block 1 and used a stepwise method of selection for all other key variables at Block 2. Thus, we performed in that way four different models: two models investigated the role of early metaphor understanding (accuracy and

specificity of mental interpretation) on the later Strange Stories (mental and physical), and two other models examined the role of early Strange Stories (mental and physical) on the later metaphor understanding (accuracy and specificity of mental interpretation).

Results

Descriptives and Preliminary Analyses

Descriptive statistics are reported in Table 3.1. Our data showed that children's scores were not at ceiling and that there was an adequate variability in scores to examine individual differences (Table 3.1).

Paired sample t-tests showed significant changes across time in Strange Stories, both mental, $t(53) = 3.56$, $p = .001$, $d = .46$, and physical stories, $t(53) = 2.23$, $p = .030$, $d = .30$, in metaphor accuracy, $t(53) = 4.42$, $p < .001$, $d = .50$, and in specificity of mental interpretation, $t(53) = 3.60$, $p = .001$, $d = .48$. Controlling for Time 1 vocabulary, developmental differences in Strange Stories performance fell below the significant level, $F(1,52) = .37$, $p = .547$, $\eta_p^2 = .007$, and $F(1,52) = 1.35$, $p = .251$, $\eta_p^2 = .025$, for mental and physical stories, respectively. Conversely, the growth in metaphor accuracy, $F(1,52) = 7.48$, $p = .009$, $\eta_p^2 = .126$, and in specificity of mental interpretation, $F(1,52) = 7.03$, $p = .011$, $\eta_p^2 = .119$, were still significant even after controlling for vocabulary.

Table 3.1. Descriptive statistics at Time 1 and Time 2.

	T1		T2		
	Mean (*SD*)	Actual Range	Mean (*SD*)	Actual Range	Possible Range
Age in years	8.86 (.23)	8.50 − 9.33	9.36 (.23)	9.00 − 9.83	−
Vocabulary	26.09 (4.27)	9 − 30	26.95 (3.34)	12 − 30	0 − 30
Mental Strange Stories	8.77 (2.33)	4 − 14	9.79 (2.08)	5 − 14	0 − 14
Physical Strange Stories	5.10 (2.10)	2 − 9	5.76 (2.26)	1 − 9	0 − 12
Metaphor accuracy	16.17 (4.25)	0 − 22	18.07 (3.63)	0 − 23	0 − 24
Physical Metaphor interpretation	6.69 (1.23)	4 − 11	6.91 (1.35)	5 − 11	0 − 18
Mental Metaphor interpretation	11.19 (2.64)	5 − 17	12.59 (2.49)	6 − 18	0 − 18
Specificity of mental interpretation	4.50 (2.38)	0 − 10	5.69 (2.58)	–2 − 10	–18 − 18

Notes. T1 = Time 1; T2 = Time 2.

Crucially, all the study variables were significantly stable across time, $r_s \geq .50$, even after controlling for vocabulary, $r_s \geq .43$ (Table 3.2).

In table 3.2 are reported the correlations and partial correlations between the key variables. Within each time point, metaphor accuracy correlated with mental and physical stories of the Strange Stories task, even after controlling for vocabulary. Differently, focusing on the specificity of mental interpretation of metaphors, we found that, within each time-point, the specificity of mental interpretation was significantly associated with mental, but not physical, stories of the Strange Stories task. Results did not change after controlling for vocabulary.

In summary, our preliminary results were the following: (1) Strange Stories performance did not improve from Time 1 to Time 2 after controlling for vocabulary; (2) metaphor accuracy and specificity of mental interpretation improved from Time 1 to Time 2 even after controlling for vocabulary; (3) all key variables were stable across time points; (4) metaphor accuracy was significantly associated with both mental and physical stories of the Strange Stories task within time points; (5) specificity of metaphor interpretation was significantly associated with mental, but not physical, stories of the Strange Stories task within time points.

Table 3.2. Correlations and Partial Correlations (below the diagonal) Between ToM, pragmatics, and vocabulary.

	T1				T2				
	M_SS	P_SS	Met_Acc	Met_Spe	Voc	M_SS	P_SS	Met_Acc	Met_Spe
T1									
Voc	.26^{+}	.29*	.34*	.42**	.50***	.29*	.47***	.16	.12
M_SS	–	.49***	.56***	.42**	.32*	.55***	.47***	.59***	.31*
P_SS	.45**	–	.40**	.24	.10	.28*	.50***	.50***	.19
Met_Acc	.52***	.34*	–	.46***	.51***	.47***	.56***	.69***	.29*
Met_Spe	.35*	.14	.38**	–	.46***	.47***	.32***	.35**	.53***
T2									
Voc	–	–	–	–	–	.29*	.27*	.42**	.20
M_SS	.51***	.22	.41**	.40**	–	–	.39**	.54***	.52***
P_SS	.41**	.43**	.48***	.15	–	.34*	–	.46**	.23
Met_Acc	.57***	.48***	.68***	.32*	–	.48***	.39**	–	.38**
Met_Spe	.29*	.16	.26^{+}	.53***	–	.49***	19	.33*	–

Notes. T1 = Time 1. T2 = Time 2. Voc = Vocabulary. M_SS = Mental Strange Stories. P_SS = Physical Strange Stories. Met_Acc = Metaphor accuracy. Met_Spe = Specificity of mental interpretation. *Scores.* *** $p < .001$ ** $p < .01$. * $p < .05$ $^{+}$ $p < .08$. Partial correlations control for concurrent vocabulary within−time, and for baseline vocabulary across−time.

Longitudinal associations between ToM and metaphor understanding

The longitudinal design allowed us to investigate not only the associations between ToM and metaphor understanding within but also across time points.

At first, we examined the relationships between early metaphor understanding and later Strange Stories. Results showed that metaphor accuracy at Time 1 significantly correlated with the performance at Time 2 in mental and physical stories of the Strange Stories task, even after controlling for vocabulary (see Table 3.2). In addition, we found that specificity of mental interpretation at Time 1 significantly correlated with the performance at Time 2 in mental and physical stories of the Strange Stories task. However, only the correlation between Time 1 specificity of mental interpretation and Time 2 mental stories of the Strange Stories task remained significant after controlling for vocabulary.

Hierarchical regression analyses, that controlled for Time 1 vocabulary and for the stability of the dependent variable, allowed us to examine in more detail the predictive effects of metaphor understanding on later ToM. Results showed that the performance at Time 2 in physical stories of the Strange Stories task was significantly predicted by metaphor accuracy at Time 1 (Table 3.3), but not by specificity of mental interpretation at Time 1, $t(53) = -.26$, $p = .795$. Differently, the performance at Time 2 in mental stories of the Strange Stories task was significantly predicted by specificity of mental interpretation at Time 1 (Table 3.3), but not by metaphor accuracy at Time 1, $t(53) = .96$, $p = .343$.

Table 3.3. Hierarchical multiple regressions on T2 mental Strange Stories and T2 physical Strange Stories.

	Dependent variables							
	T2 mental Strange Stories				T2 physical Strange Stories			
	B	SE	β	95% CI	B	SE	β	95% CI
Step 1								
T1 vocabulary	.08	.06	.16	[−0.04, 6.99]	.19	.06	.35**	[0.06, 0.31]
T1 mental Strange Stories	.46	.11	.51***	[0.24, 0.67]	–	–	–	–
T1 physical Strange Stories	–	–	–	–	.43	.13	.40**	[0.18, 0.68]
T1 specificity of mental interpretation	–	–	–	–	–	–	–	–
T1 metaphor accuracy	–	–	–	–	–	–	–	–
Step 2								
T1 vocabulary	.03	.06	.07	[−0.09, 0.15]	.14	.06	.27*	[0.02, 0.26]
T1 mental Strange Stories	.38	.11	.42**	[0.16, 0.60]	–	–	–	–
T1 physical Strange Stories	–	–	–	–	.30	.12	.28*	[0.05, 0.55]
T1 specificity of mental interpretation	.24	.12	.27*	[0.01, 0.47]	–	–	–	–
T1 metaphor accuracy	–	–	–	–	.19	.07	.35**	[0.06, 0.31]
	Adjusted R^2 = .30*** for step 1				Adjusted R^2 = .34** for step 1			
	Adjusted ΔR^2 = .05* for step 2				Adjusted ΔR^2 = .10** for step 2			

Notes. T1 = Time 1. T2 = Time 2. B = Unstandardized Beta. SE = Standard Error. β = Standardized Beta. CI = Confidence Interval. *** $p < .001$. ** $p < .01$. * $p < .05$.

We then examined the associations between early Strange Stories and later metaphor understanding. Results showed that performance at Time 1 in mental stories of the Strange Stories task significantly correlated with metaphor accuracy at Time 2 and specificity of mental interpretation at Time 2, even after controlling for vocabulary (Table 3.2). On the contrary, performance at Time 1 in physical stories of the Strange Stories task significantly related with metaphor accuracy at Time 2, but not with specificity of mental interpretation at Time 2 (Table 3.2). Results did not change after controlling for vocabulary. Hierarchical regression analyses, that controlled for vocabulary at Time 1 and for the stability of the dependent variable, allowed us to examine in more detail the predictive effects of ToM on later metaphor understanding. Results showed that metaphor accuracy at Time 2 was significantly predicted by physical stories of the Strange Stories task at Time 1 (Table 3.4).

Table 3.4. Hierarchical multiple regressions on T2 metaphor accuracy, and T2 specificity of mental interpretation.

	Dependent variables							
	T2 metaphor accuracy				T2 specificity of mental interpretation			
	B	SE	β	95% CI	B	SE	β	95% CI
Step 1								
T1 vocabulary	−.07	.09	−.08	[−.26, .11]	−.07	.08	−.12	[−.23, .09]
T1 mental Strange Stories	–	–	–	–	–	–	–	–
T1 physical Strange Stories	–	–	–	–	–	–	–	–
T1 specificity of mental interpretation	–	–	–	–	.63	.14	.58***	[.34, .91]
T1 metaphor accuracy	.61	.09	.72***	[.43, .80]	–	–	–	–
Step 2								
T1 vocabulary	−.11	.09	−.13	[−.29, .06]				
T1 mental Strange Stories	–	–	–	–				
T1 physical Strange Stories	.51	.18	.29**	[.14, .88]				
T1 specificity of mental interpretation	–	–	–	–				
T1 metaphor accuracy	.52	.09	.61***	[.34, .71]				
	Adjusted R^2 = .46*** for step 1 Adjusted ΔR^2 = .07** for step 2				Adjusted R^2 = .29*** for step 1 –			

Notes. T1 = Time 1. T2 = Time 2. B = Unstandardized Beta. SE = Standard Error. β = Standardized Beta. CI = Confidence Interval. *** $p < .001$. ** $p < .01$.

Interestingly, once we performed an analysis in which we forced the entry of mental stories of the Strange Stories task at Time 1 before the physical stories at Time 1, we found that mental stories Time 1 had a significant role in predicting later metaphor accuracy, $t(53) = 2.68$, $p = .010$, but this effect fell below the statistical significance level after the entry of physical stories of the Strange Stories task at Time 1 in the model, $t(53) = 1.90$, $p = .064$. On the contrary, specificity of mental interpretation at Time 2 was significantly predicted only by its scores at Time 1 (Table 3.4). Thus, we found no significant effect either of mental stories of the Strange Stories task at Time 1, $t(53) = .98$, $p = .330$, or of physical stories of the Strange Stories task at Time 1, $t(53) = .74$, $p = .465$, in predicting specificity of mental interpretation at Time 2.

In summary, our results across time, controlling for vocabulary and stability, were the following: (1) the performance at physical stories of the Strange Stories task at Time 2 was significantly predicted by metaphor accuracy at Time 1; (2) the performance at mental stories of the Strange Stories task at Time 2 was significantly predicted by specificity of mental interpretation at Time 1; (3) metaphor accuracy at Time 2 was significantly predicted by physical stories of the Strange Stories task at Time 1; (4) specificity of mental interpretation at Time 2 was significantly predicted only by its own stability. This pattern of results suggests a bidirectional relationship between metaphor accuracy and Strange Stories, especially the physical stories, and a unidirectional relationship between specificity of mental interpretation and mental, but not physical, stories of the

Strange Stories task in that specificity of mental interpretation at Time 1 predict later performance at mental stories of the Strange Stories task.

Discussion

The main aim of the present longitudinal study was to examine the directionality of the relationships between individual differences in metaphor understanding and ToM in a sample of typically developing middle-aged children. The present study distinguished between metaphor accuracy and specificity of mental interpretation, and between the ability to make inferences about mental (as indexed by mental stories of the Strange Stories task) and physical causes (as indexed by physical stories of the Strange Stories task). Having considered these two aspects of the Strange Stories let us clarify whether metaphor understanding (accuracy and/or specificity of mental interpretation) is specifically related to ToM rather than to a more general inferential skill and enabled us to identify specific rather than general associations.

Before discussing the results about the direction of the relationships between these two constructs, we present several preliminary results concerning (a) the developmental changes, (b) the stability, and (c) the concurrent relationships across study variables.

Regarding the developmental changes (a), while metaphor understanding significantly improved over the 6 months (independently of children's vocabulary), the improvement in ToM skills fell below statistical significance after controlling for children's vocabulary. The latter result is in contrast with that

of Lecce and colleagues that found a significant improvement in ToM abilities over periods of 6 months in children of the same age as the present sample (Lecce et al., 2017). However, Lecce and colleagues did not control for verbal abilities, leaving open the possibility that ToM changes in the short term may be explained by more general changes in language ability.

Regarding the key-variables stability (b), as expected, all measures were stable over time, even after controlling for children's vocabulary, suggesting that variations in scores on our tasks reflected a genuine variability in the underlying construct.

Focusing on the extended version of the PMM task and taking together these two preliminary results, we can conclude that the extended version of the PMM task is a sound measure for detecting individual differences in metaphor interpretation and metaphor comprehension during middle childhood.

Regarding the concurrent associations between metaphor understanding and ToM (c), we found concurrent relationships between children's ability to understand the meaning conveyed by a metaphor (i.e., metaphor accuracy) and their ability to make inferences both about the mental (i.e., ToM) and physical causes (i.e., general inferential skill) of others' social behaviors. These links existed independently from children's vocabulary. Differently, specificity of mental interpretation was specifically linked with the concurrent ability to make inferences about the mental (i.e., ToM), but not physical states (i.e., general inferential skill), even after controlling for children's vocabulary. The present results are in line with a previous study showing a link between ToM and

metaphor comprehension in middle childhood (Whyte & Nelson, 2015) and with results from Chapter 2, that reported a specific link between mental metaphor interpretation and ToM.

Moving to the main aim, we first discuss the longitudinal relationships between *metaphor accuracy* and both ToM and general inferential skill. Interestingly, our results showed longitudinal relationships between children's metaphor comprehension and both ToM and general inferential skill, even after controlling for children's vocabulary. However, when we also controlled (for the first time in the literature) for the stability of individual differences, we found that only general inferential skill (i.e., the ability to make inferences about physical states), but not ToM (i.e., the ability to make inferences about mental states), was bi-directionally linked with metaphor comprehension. Interestingly, we also found that early ToM significantly predicted later metaphor comprehension, but that this relationship fell below the statistical significance level when we controlled for the more general inferential ability. Overall, this pattern of results suggests that the ability to make inferences, not specifically about mental states, is strongly implicated in the development of metaphor comprehension and vice versa. In addition, the present results led us to the conclusion that the association between accuracy in metaphor understanding and ToM could be, at least in part, explained by a general inferential ability. This is a novel finding in the literature since no studies already investigated the association between general inferential ability and metaphor understanding in typically developing middle-aged children. Evidence in line with this finding comes from studies reporting a relationship

between reading comprehension, which includes inferential questions, and both ToM and metaphor understanding in typically developing middle-aged children (Cantin et al., 2016; Seigneuric et al., 2016). Our findings are also supported by evidence coming from the literature on adult clinical populations that found an association between pragmatic abilities and non-mental-specific inferential skills. For example, Montemurro and colleagues (2019) examined a sample of individuals with Parkinson's disease and showed that their pragmatic skills were related to the ability to draw inferences, both about mental aspects and about physical causality, as assessed with the Story Empathy Task. On the same line, Martin & McDonald (2005) showed that the impairment of individuals with traumatic brain injury in understanding irony was mainly due to a more general inferential reasoning capacity, rather than to ToM skills. We found a similar pattern of results in typically developing middle-aged children, suggesting that the comprehension of metaphors, once we control for verbal skills and task stability, is indeed not linked to ToM specifically, but mainly to a broader inferential reasoning capacity.

We then discuss the longitudinal relationships between *specificity of mental interpretation* of metaphor and both ToM and general inferential skill. Our results showed significant bidirectional longitudinal associations between the specificity of mental interpretation of metaphors and ToM. Crucially, these relationships were specific in that: (a) they were independent of vocabulary, and (b) they did not extend to children's ability to make inferences about physical states. However, when we also controlled for the stability of individual

differences, we found that the specificity of mental interpretation of metaphor at Time 1 predicted ToM at Time 2, but ToM at Time 1 did not predict specificity of mental interpretation at Time 2. This result is totally new in the literature and should be commented on, considering the two directions apart. Regarding the first result showing that ToM at Time 1 did not predict specificity of mental interpretation at Time 2, it is important to note that we measure ToM via an advanced task, namely the Strange Stories task. While the more classical tasks of ToM assessed more basic abilities (such as the ability to understand conflicting representations involved in knowledge-ignorance or false-belief), the Strange Stories task assessed more complex ToM skills (such as the ability to make inferences about a variety of others' mental states in complex social scenarios). It is therefore possible that considering this more advanced aspect of ToM does not allow us to detect the effect of ToM on later specificity of mental interpretation. Future studies should also assess ToM using classical tasks to shed light on this topic. To comment on the second result showing that the specificity of mental interpretation of metaphor at Time 1 predicted ToM at Time 2 we should consider evidence reporting a link between ToM and mental state lexicon. Indeed, specificity of mental interpretation reflects a tendency to spontaneously prefer a mental, rather than a physical, interpretation when possible and appropriate. This tendency can be viewed as a preferred orientation towards mental states and as a mentalistic style that, in turn, increases the understanding of the mental states behind social behavior, namely ToM. In line with this view, a robust number of studies brings supports to our findings showing that those children who score

higher on classic ToM tasks refer more frequently to mental states in spontaneous conversations (Brown, Donelan-McCall, & Dunn, 1996; Hughes, Lecce, & Wilson, 2007). This relation was found also in studies on school-aged children that examined the frequency of mental state utterances used during non-interactional tasks (Lecce et al., 2010; Ornaghi & Grazzani, 2013). Interestingly, longitudinal studies in pre-schoolers children suggested that even when controlling for age and language, mental state lexicon predicted later ToM but not vice versa (Hughes & Dunn, 1998; Ruffman, Slade, & Crowe, 2002).

Discussing our findings in light of the three pattern of results that we have hypothesized, we found no support either for the first hypothesis (i.e., ToM has a role in the development of metaphorical skills) or for the second hypothesis (i.e., metaphorical understanding drives the development of ToM). Our results seem rather in favor of the third hypothesis, according to which the direction of the developmental associations between ToM and metaphor understanding varies depending on the pragmatic aspect considered and is in part explained by the role of general inferential ability. Indeed, when we assessed the ability to understand the meaning conveyed by a metaphor (i.e., metaphor accuracy), the relationships between ToM and metaphor understanding were driven by a general inferential ability. Given the bidirectional nature of these relationships, we support a view in which inferential and metaphor comprehension abilities develop side by side in a mutually supportive way. Conversely, when we assessed the ability to interpret mentally mental, but not physical, metaphors, metaphorical skills predict later ToM abilities.

Caveats and Conclusions

The present study is the first that examined the developmental associations between ToM and metaphor understanding in typically developing children using a fine-grained analysis that distinguished between inferences about mental vs. physical states and between comprehension vs. interpretation of mental and physical metaphors. Theoretically, both inferences about mental and physical states (measured via mental and physical storied from the Strange Stories) and inferences about the metaphorical meanings are forms of reasoning by which we derive reasonable consequences from premises. In line with this, our results showed that children's ability to comprehend metaphors was closely linked with broader inferential ability, in a bidirectional way. When we considered another aspect of metaphor understanding, more related to the kind of interpretation given by the children, then we observed that the ability to appropriately interpret mental metaphors and to verbalize such interpretations was predictive of later ToM development. In other words, being able to capture the mental aspects in a metaphorical communication might promote the ability to infer mental states underlying social behavior and, more generally, to reason about others' mental states.

The present results contribute to clarify the directionality of the relationships between ToM and metaphor understanding in middle childhood. However, we have just scratched the surface of the complex relationships between pragmatics, metaphors, and ToM. Indeed, the present study is not exempt from limitations.

First, larger samples and longer periods of time are needed to confirm our results.

Secondly, regarding the measures, using single tasks for assessing ToM and for metaphor understanding prevent us to generalize the present results to the broader constructs of ToM and metaphor understanding. Future studies should expand the scope of the present research by including other measures. For example, it could be interesting to examine the link between ToM and metaphor understanding focusing on non-verbal ToM and on ToM sub-components, (e.g. trough Triangle task; Castelli, Happé, Frith, & Frith, 2000, and hindsight task; Bernstein, Atance, Loftus, & Meltzoff, 2004). In addition, future studies should also assess metaphor production. In this respect, Clark (2019) claimed that pretend play and perspective-taking, which are known to be proximal of ToM, are precursors of similes and metaphors production as they reflect the ability to consider more than one mental representation at the time. According to this view, while metaphor comprehension is mutually related, in a bidirectional way, with inferences in general, metaphor production should be predicted by ToM and not vice versa.

Despite these limitations, the present study has, in addition to theoretical implications, a practical impact leading to the shaping of training programs for promoting communicative and socio-cognitive skills in children.

Chapter 4 – Theory of Mind and Metaphor Understanding: a training study in middle childhood

Abstract

The main aim was to investigate the existence of causal relationships between ToM and metaphor understanding during middle childhood. To this aim, we conducted a training study on 55 typically developing children (age = 8;9 – 9;10 years) assigned to two training conditions: ToM or MetaCom. Both training programs were made up of four sessions involving group conversations. We trained children's ability to make context-sensitive inferences about mental states (ToM) or metaphorical meanings (MetaCom). We measured, at baseline and after the end of the training programs ToM (via Strange Stories and Triangles) and metaphor understanding (via a hard version of the Physical and Mental Metaphors [PMM] task and the referential task). At baseline, children were also tested for children's working memory, reading comprehension, and verbal abilities. Preliminary results on the soundness of the hard version of the PMM task showed that it is an adequate measure for capturing individual differences in metaphor understanding, during middle childhood. Main results showed the two training groups were equal in all study variables at baseline. We found that both training programs were effective in enhancing the trained variables. Crucially, while children in ToM group improved in their ability to understand context-sensitive, but not minimal-context, metaphors, children in the MetaCom group enhanced their ToM skills. No training was effective in enhancing children's specificity of mental interpretation of metaphors. This pattern of

results suggests that while training ToM enhances the ability to make context-sensitive, but not minimal-context, inferences about metaphorical meanings, training metaphor understanding improves children's ability to attribute and infer others' mental states.

Introduction

As we have already anticipated in previous chapters, till now the literature about the relationships between ToM and metaphor understanding in middle-aged typically developing children is scarce. Most of the studies that reported relationships between metaphorical skills and ToM were cross-sectional (Whyte & Nelson, 2015; Study 2).

In addition, findings from Study 3 of the present dissertation (see Chapter 3) tried to shed light on the directionality of these links, finding no support either to the hypothesis that ToM has a role in the development of metaphorical skills or to the hypothesis that metaphorical understanding drives the development of ToM. Results from Chapter 3 are in favor of the hypothesis according to which the direction of the developmental associations between ToM and metaphor understanding varies depending on the pragmatic aspect considered, namely accuracy vs. specificity of mental interpretation of metaphors. Indeed, while inferential skills and metaphor comprehension abilities develop side by side in a mutually supportive way, the ability to interpret mentally mental, but not physical, metaphors, predicts later ToM, but not vice versa.

However, these results come from a longitudinal design study that is able to demonstrate the genuine developmental relationships but not the causal nature of these relationships (Bradley & Bryant, 1983).

Starting from these premises, the present study tried to fill this gap in the literature by investigating the causality of the relationships between ToM and metaphor understanding in 9-year-old typically developing children adopting a training design.

The literature on ToM training for typically developing children is quite large and mainly focused on the importance of conversations. For example, two studies on this topic showed that involving 3- and 4-year-olds in conversations about beliefs, desires, and perceptions lead to higher improvements in their false belief understanding compared to children in a control group (Appleton & Reddy, 1996; Slaughter & Gopnik, 1996). Interestingly, other studies highlighted the important role of feedbacks and explanations to children's answers for enhancing their ToM skills (Clements, Rustin, & McCallum, 2000; Melot & Angeard, 2003). More recently, Lecce and colleagues (2014) tried to put this evidence together in order to develop an effective training program for middle childhood. In detail, this training focused on the use of (a) children's conversations, and (b) feedbacks and explanations. In addition, they decided to focus on conversations about mental states capitalizing on the idea that the frequency, the quality, and the content of conversations predict later ToM skills (Appleton & Reddy, 1996; Ensor & Hughes, 2008; Ornaghi, Brockmeier, & Gavazzi, 2011; Peterson & Slaughter, 2003). This training program is widely used in the literature on middle childhood

and was found to be effective in enhancing children's ToM (Bianco et al., 2016; Lecce & Bianco, 2018; Lecce, Bianco, Devine, Hughes, & Banerjee, 2014). For example, Lecce and colleagues (2014) found that 9- to 10-year-old children that underwent this ToM training improved significantly more, both in the short and in the long term, in their ToM skills compared to children that attended an active control training. Interestingly, this gain in ToM skills was independent of any change in executive functions. Bianco and colleagues (2016) found similar results showing that the ToM training was effective in enhancing ToM, measured via a practiced task and a transfer task, in children aged 9 to 10 years.

The literature on training metaphor understanding mainly focused on children with developmental disorders, in particular with ASD children (Mashal & Kasirer, 2011; Melogno et al., 2017; Persicke, Tarbox, Ranick, & St. Clair, 2012). Only two studies developed training programs for enhancing metaphor comprehension in typically developing children. One of these studies aimed to train children's metaphor comprehension enhancing their ability to identify the metaphorical link between the topic and the vehicle of metaphors (Białecka-Pikul, 2010). However, this training was quite implicit and, thus, was not effective in enhancing metaphor understanding in 4- to 5-year-old children (Białecka-Pikul, 2010). More recently, Tonini and colleagues (2020) designed a Metaphor Comprehension (MetaCom) training program that is grounded on a strong theoretical framework, namely the relevance-theoretic lexical account of metaphor (Wilson & Carston, 2007). Compared to the previous training, the one by Tonini and colleagues was more explicit and capitalized on the idea that the

comprehension of nominal metaphors involves the context-driven adjustment of the lexically encoded concept, and that a non-literal interpretation is derived through a series of pragmatic inferences about relevant properties of the vehicle. Like the training of Lecce and colleagues (2014) they focused on the use of (a) children's conversations, and (b) feedbacks and explanations. Authors found that 9-year-old children that underwent the MetaCom training improved more compared to children in an active control group in metaphor understanding, measured via a practiced task and a transfer task (Tonini et al., 2020).

The Present Study

The main aim of the present study was to investigate the causal relationships between ToM and metaphor understanding in 9-year-old typically developing children. To reach this goal, we adopted a training design and compared the effects of two training programs, one targeting ToM (taken from Lecce et al., 2014) and the other targeting metaphor comprehension (MetaCom; taken from Tonini et al., 2020). Before investigating the main aim, we checked that the group of children that underwent the ToM training and the group of children that experienced the MetaCom training were equal for children's receptive language, grammar, reading comprehension, and working memory, which are known be related to both metaphors (Carriedo et al., 2016; Norbury, 2005; Seigneuric et al., 2016) and ToM (Cantin et al., 2016; Ebert, 2020; Hughes, 1998; Lecce et al., 2017).

The two intervention programs (ToM and MetaCom) were conversation-based training and were matched in structure and length (see Procedure section for further details). Indeed, both training programs were composed of 4 sessions of about 50 minutes each. Each session was composed of stories and language exercises in a written form. After each story children have to make inferences about the story. After each story and each exercise children were encouraged to take part in group conversations about their answers. The only difference between the ToM and the MetaCom training is the content of the stories and the exercises.

The ToM training was the one developed by Lecce and colleagues (2014). During ToM training children were taught to attribute and understand others' mental states and their dynamic nature. The stories and the exercises of the ToM training required children to make context-sensitive inferences about the main character's mental state in order to make sense of what is happening in a specific situation (see Appendix B for an example). Stories in the ToM condition were two misunderstanding, two faux pas, two double-bluff, and two persuasion stories based on the Strange Stories task (White et al., 2009). After each story children were asked to infer the mental states underlying the main character's behavior. Language exercises in the ToM condition were based on the metacognitive language task by Olson and colleagues (Olson, Antonietti, Liverta-Sempio, & Marchetti, 2006). In those exercises, children were asked to select the correct synonyms for mental state verbs.

Conversely, the MetaCom training was an intervention recently developed by Tonini and colleagues (2020). During MetaCom training children were taught

to select the relevant properties of metaphorically used concepts, as well as to use the expressions in new contexts. The stories and the exercises of the MetaCom training required children to make context-sensitive inferences about the meaning of nominal metaphors (see Appendix C for an example). Stories in the MetaCom condition represented several different social situations in which nominal metaphors in the form "X is Y" were embedded. After each story children were asked to adjust the lexical concepts by selecting the relevant properties and inferring the metaphor meaning using the context. Interestingly, three metaphorical stories included in the training conveyed a mental meaning of the metaphor (e.g., the metaphor "The teacher is a lantern" meaning "He clarifies children's doubts"). Language exercises in the MetaCom condition were built to strengthening the comprehension of the metaphor meaning, promoting the generalization to other contexts and the use of metaphors. In those exercises, children were asked to create new metaphors and to tell a story containing the learned metaphor.

As we have already mentioned, the efficacy of the MetaCom and the ToM training compared to active control training were already shown in previous studies (Bianco & Lecce, 2016; Bianco et al., 2016; Lecce & Bianco, 2018; Lecce et al., 2014; Tonini et al., 2020).

In order to investigate the existence of causal relationships between ToM and metaphor understanding in 9-year-old typically developing children, we tested the efficacy of the two training (ToM and MetaCom) on children's ToM and metaphor understanding.

Specifically, we focused on the effect of the ToM training on metaphor understanding measured via two tasks: the first task was a modified version of the PMM task (see Measures section) and it was used to measure children's ability to understand nominal metaphors, namely the type of metaphor directly trained with the MetaCom training. We considered this task as the MetaCom practice task. The second task was the referential task (see Measures section) and it was used to measure children's ability to understand referential metaphors, namely the type of metaphor different from the one trained with the MetaCom training. We considered this task as the MetaCom transfer task. Since ToM training targeted children's ability to infer others' mental states and given results from Study 3 (see Chapter 3) showing that inferential abilities, not limited to inferences about mental states, predicted later metaphor understanding, we expected that ToM would extend its effects on the ability to understand metaphors. However, we did not have predictions regarding possible different effects of the ToM training on the MetaCom practice task and the MetaCom transfer task.

In addition, we focused on the effect of the MetaCom training of ToM measured via two tasks: the first task was the Strange Stories task (see Measures section) and it was used to measure children's ability to understand others' mental states within different social scenarios similar to the ones used during the ToM training. This task measured the ToM that was directly trained with the ToM training. Thus, we considered the Strange Stories task as the ToM practice task. The second task was the animated triangles task (see Measures section) and it was used to measure children's ToM via a task that was different in structure (humans

vs. geometric shapes) and modality (verbal static vs. visual dynamic) from the stories used in the ToM training. Thus, we considered this task as the ToM transfer task. Since MetaCom training focused on the pragmatic inferential mechanisms underlying the comprehension of metaphors, including mental ones, and given results from Study 3 (see Chapter 3) showing that metaphor interpretation predicted later ToM, we expected that the MetaCoM training would extend its effects on ToM. However, we did not have predictions regarding possible different effects of the MetaCoM training on the ToM practice task and the ToM transfer task.

Materials and Method

Participants

66 children took part in the present study. 31 typically developing children (16 M, Pre-training *M*age = 9;3, *SD* = 0;4, age-range = 8;7 – 10;3) were recruited from four classes in three primary schools in Northern Italy. After the pre-training assessment, children underwent a ToM training condition. Additionally, 35 typically developing children (15 M, Pre-training *M*age = 9;3, *SD* = 0;4, age-range = 8;9 – 9;10) were recruited within the MetaCom study (Tonini et al., 2020) in which, after the pre-training assessment, children underwent a MetaCom training condition.

Exclusion criteria were (a) having a diagnosis of specific learning disorder or developmental delay, (b) not speaking Italian from birth, (c) missing the pre- or post-training assessment, and (d) scoring at ceiling in the trained tasks, at the pre-

training assessment. Thus, the final sample consisted in 55 typically developing children (26 M, Pre-training *M*age = 9 years; 3 months, *SD* = 0;3, age-range = 8;9 – 9;10). The ToM group consisted of 23 children (13 M, Pre-training *M*age = 9;2, *SD* = 0;3, age-range = 8;9 – 9;9), and the MetaCom group consisted of 32 children (13 M, Pre-training *M*age = 9;3, SD = 0;3, age-range = 8;9 – 9;10).

Procedure

Before the assessment, parental written consent was collected for all participants. We measured, before the beginning and after the end of training programs, children's ToM, and metaphor understanding. At the pre-training assessment, children were also tested for verbal abilities, reading comprehension, and working memory.

After the end of the pre-training assessment, children took part in one of the two training programs (ToM or MetaCom). Each training program was made up of four sessions involving group conversations. Each intervention session lasted about 50 minutes and was made up of two stories and two exercises. After each story and each exercise children were asked to write down their answers individually and, then, were encouraged to take part in group conversations about their answers. This study has been approved by the local University Ethical Committee.

Measures

Verbal abilities. Verbal abilities were evaluated via the Test for Reception of Grammar–version 2 (TROG–2; Bishop, 2003) and the Peabody Picture Vocabulary Test–Revised (PPVT–R; Dunn & Dunn, 1981).

TROG–2 (Bishop, 2003; Suraniti, Ferri, & Neri, 2009) evaluated children's comprehension of grammar structure. It includes 80 items divided into 20 blocks, each addressing a specific grammatical structure. We administered to children four blocks regarding the understanding of subject relative clause (block G), reversible passive (block K), singular and plural inflection (block R), and object relative clause (block S).

After two example items that were considered with the experimenter, children were asked to match a sentence to the corresponding picture on a printed booklet. The total score could range from 0 to 16.

PPVT–R (Dunn & Dunn, 1981; Stella, Pizzioli, & Tressoldi, 2000) evaluated children's receptive language. It includes 175 items. The task requires that for each child a starting point of the test is determined based on his/her age and going backward until the child does not provide 8 consecutive correct answers. The task ends after six errors out of eight consecutive answers. Of the total amount of items, we selected 65 items, starting 20 items before the item identified as the starting point for age 9 (i.e., item 70) in order to administer this test collectively. After three example items considered with the experimenter, children were asked to match a word to the corresponding picture on a printed

booklet. The final score was calculated according to the test manual and could range from 50 to 115.

Reading comprehension. Reading comprehension was evaluated through the memory and transfer standardized reading comprehension battery (MT; Cornoldi & Colpo, 1998). Children were asked to answer 10 multiple-choice questions after having read a brief narrative. Answering correctly to these questions required inferential reasoning skills since the questions didn't probe literal information. The total score could range from 0 to 10.

Working memory. Working memory was measured via the Backward Digit Span task taken from the Italian version of WISC-R (Orsini, 1993). Children had to recall in reverse order seven series of digit sequences just listened. The length of the series ranged from two to eight digits, presented ordinately with an increasing level of difficulty. The total working memory score could range from 0 to 7.

Theory of Mind. Theory of Mind was assessed through the Strange Stories task (Happé, 1994) and the animated triangles tasks (Castelli et al., 2000).

Strange Stories task (Happé, 1994) assessed children's ability to interpret the main character's utterances by making inferences about the character's behavior. We administered seven mental stories (two double bluffs, two misunderstandings, two persuasions, and one white lie) and five physical/control

stories. Mental stories require children to make context-appropriate inferences about the character's mental states, whereas physical stories require children to make context-appropriate inferences about the physical cause of the character's behavior. After reading each story, children were asked to answer an open-ended question in which they had to explain the reasons underlying the main character's behavior in a written form. Following the scoring guidelines (White et al., 2009) 0 points were assigned for incorrect and "Don't know" answers, 1 point for partially correct and implicit answers, and 2 points for full and explicit answers. The total ToM score (given by the performance at the mental stories of the Strange Stories task) could range from 0 to 14 while the total inferential score (given by the performance at the physical stories of the Strange Stories task) could range from 0 to 10. We established the inter-rater agreement using Cohen's kappa on the 25% of the responses, reporting an almost perfect agreement at both time points (k = .80 at Pre-training and k = .89 at Post-training). The total inferential score was used to check for groups equivalence at pre-training assessment while the ToM score was considered as the primary outcome measure of the ToM training since it assessed practiced ToM ability (i.e., the ability to infer the main character's mental state and, thus, to make sense of what is happening in a specific situation).

Animated triangles task (Castelli et al., 2000) assessed children's ability to attribute mental states to two triangles of different sizes moving on a screen. We administered three short animations involving surprise, deceive and mocking. Children, after having watched each video-clip, had to write down what happened. Each animation was scored according to the level of intentionality attributed to

triangles. The score for each item ranged from 0 (no deliberate action) to 5 (deliberate action aimed at manipulating others' mental state). The total score could range from 0 to 15. We established the inter-rater agreement using Cohen's kappa on the 25% of the responses, reporting almost perfect agreement at Pre-training and a substantial agreement at Post-training (k = .82 at Pre-training and k = .68 at Post-training). Since this task measured a ToM ability (i.e., the tendency to attribute mental states) different from the trained one we considered it as a transfer measure of the ToM training.

Metaphor Understanding. Metaphor understanding was evaluated via the Physical and Mental Metaphors (PMM) task and the referential task (Noveck et al., 2001).

PMM task assessed children's ability to understand nominal metaphors in a minimal context (see previous chapters). We used a modified and hard version of the 12-items PMM task (see Chapter 1). Since we want to observe an improvement after the training, we exclude the metaphors that in Study 1 and 3 scored at ceiling in accuracy and in interpretation in more than 70% of children at both time points. Thus, we replaced the metaphors (a) "Il calciatore è una freccia" (literal translation "The footballer is an arrow"), (b) "Mio fratello è un grattacielo" ("My brother is a skyscraper"), and (c) "La mamma è un cioccolatino" ("Mummy is a candy") with the metaphors (a) "Quella sposa è una nuvola" ("That bride is a cloud"), (b) "Quel pugile è un panda" ("That boxer is a panda"), and (c) "le mamme sono agende" ("Moms are agendas"). We found that the two new sets

were matched for the frequency of target word, familiarity and aptness of metaphors, $t(10) = 1.13$, $p = .286$, $t(10) = -1.33$, $p = .213$, $t(10) = -1.25$, $p = .240$, respectively. In addition, we checked the distinction between physical and mental metaphors, finding that metaphors in the physical set scored significantly higher on physical attributes compared to metaphors in the mental set, $t(10) = 5.72$, $p < .001$, and that metaphors in the mental set scored significantly higher on mental attributes compared to metaphors in the physical set, $t(10) = -20.77$, $p < .001$. As for the 12-items PMM task (see Chapter 1), within each set, the difference between the two interpretations was significantly bigger for metaphors in the physical set compared to metaphors in the mental set, $t(10) = 3.95$, $p = .003$. For a full description of the ratings and the comparison between the two sets of the PMM hard version see Tables 1.1b and Table 4.1.

Table 4.1. Characteristics of the new items of the hard version of the Physical and Mental Metaphors task

Metaphor	English literal translation	Set	Familiarity	Physical score	Mental score	Aptness	Frequency
Quella sposa è una nuvola	That bride is a cloud	Physical	3.68	5.58	1.96	4.00	208.48
Quel pugile è un panda	That boxer is a panda	Physical	2.42	4.67	1.33	2.79	8.48
Le mamme sono agende	Mums are agendas	Mental	3.57	3.61	5.07	4.32	0

Notes: In the table, we report the ratings of familiarity, physical and mental characteristics, aptness, and frequency values in school-age children's language (based on Marconi et al., 1994). These ratings were obtained in the way described in Chapter 1 and from 53 subjects (40 F; *M*age = 23.91, *SD* = 2.33, age-range = 21 – 32; *M*education = 15.83, *SD* = 1.67, education range = 13 – 18).

During the PMM task, children were asked to verbally explain the meaning of each nominal metaphor. Following the scoring guidelines (see Chapter 1 and Chapter 3), metaphors were coded according to their level of accuracy and interpretation. Accuracy, defined as the ability to articulate the link between the topic and the vehicle, was coded on a 3-point scale: we attributed 0 for incorrect, literal, and "Don't know" answers; 1 for answers that were incomplete or referred to non-salient features of the metaphor topic; and 2 for answers that were complete and referred to salient features of the metaphor topic. Inter-rater agreement was established through Cohen's kappa on the 25% of the answers, finding an almost perfect agreement at both time points (k = .85 and k = .84 for Pre-training and Post-training, respectively). A total accuracy score was calculated by summing the scores obtained at each metaphor (range = 0 – 24).

Interpretation, defined as the ability to interpret metaphors as mental or physical, was attributed on a 4-point scale (see Chapter 3): we attributed 0 for "Don't know" answers; 1 for answers referred to physical attributes or actions of the metaphor topic; 2 for ambiguous answers that could be referred either to psychological or physical attributes of the metaphor topic; and 3 for answers referred to psychological attributes of the metaphor topic. Inter-rater agreement was established through Cohen's kappa on the 25% of the answers, finding an almost perfect agreement at both time points (k = .83 and k = .89 for Pre-training and Post-training, respectively). Specificity of mental interpretation (i.e., an index that reflects children's ability to interpret mentally only mental and not physical metaphors) was calculated by subtracting the interpretation score of physical

metaphors from the interpretation score of mental metaphors (range = –18−18). We considered accuracy in the PMM as the primary outcome measure (practice task) of the MetaCom training, since it assesses the comprehension of the same type of metaphor, namely nominal metaphors, as the trained one. Since the MetaCom training included stories conveying a mental meaning of the metaphor topic we also addressed if MetaCom training generalized its effect on specificity of mental interpretation.

Referential task (adapted from Noveck et al., 2001) assessed children's ability to understand a referential metaphor within a story-context. It includes 16 pairs of eight-line stories in which the first line introduced a referent and the seventh line contained a referential term that could be either metaphorical or literal (synonymic). We created two lists of 24 stories (eight metaphorical, eight literal, and eight filler stories). At the end of each story, children were asked to answer a yes-or-no question regarding the identification of the referent. For literal and metaphorical stories, the answer was always "yes", while for fillers stories the answer was always "no". Following the scoring guidelines (Noveck et al., 2001) total accuracy score was obtained by summing the number of correct answers for metaphorical and literal conditions (range = 0 – 8). The score in the literal condition was used to check the equivalence of groups at Pre-training. The score in the metaphorical condition was considered as a transfer measure of the MetaCom training since this task measured the comprehension of metaphors (i.e., referential metaphors) different from the trained ones.

Statistical analyses plan

Before investigating our main aim (i.e., the causal associations between ToM and metaphor understanding in middle childhood), we conducted preliminary analyses. Thus, we ran a series of independent-sample t-tests in order to verify that the two training groups did not differ in any variable measured at the pre-training assessment.

In the main analyses, we performed five repeated-measures ANOVAs in order to investigate the training effect on the following dependent variables: (a) accuracy in the PMM task (main outcome of the MetaCom training), (b) performance at the mental stories of Strange Stories task (main outcome of the ToM training), (c) accuracy in the Referential task (transfer of the MetaCom training), (d) intentionality score in the animated triangles task (transfer of the ToM training), and (e) specificity of mental interpretation. Time (Pre-training and Post-training) was included as the within-subjects factor and group (MetaCom and ToM) was included as the between-subjects factor.

Results

Descriptives and Preliminary Analyses

Table 4.2 shows the descriptive statistics of all study variables divided into the two groups. Our data showed that children's scores were not at ceiling and that there was an adequate variability in scores to examine individual differences (Table 4.2).

Table 4.2. Descriptives at the pre- and post-training assessment and group differences (ToM vs. MetaCom) at baseline.

	ToM (N = 23)		MetaCom (N = 32)		
	Mean (*SD*)	Range	Mean (*SD*)	Range	*p*
Pre-training assessment					
Age in months	110.61 (3.85)	105 – 117	111.69 (3.96)	105 – 108	.318
Grammar	15.22 (1.00)	13 – 16	15.41 (0.95)	13 – 16	.478
Vocabulary	93.57 (4.73)	82 – 100	93.59 (8.11)	65 – 101	.987
Working memory	2.26 (0.86)	1 – 4	2.19 (0.78)	1 – 4	.744
Reading comprehension	7.17 (2.21)	1 – 10	7.13 (1.88)	2 – 10	.930
Physical Strange Stories	4.61 (2.08)	1 – 9	4.63 (1.43)	2 – 7	.973
Mental Strange Stories	7.91 (2.66)	3 – 13	8.19 (2.67)	4 – 13	.708
Animated triangles	10.41 (2.62)	6 – 15	10.97 (2.39)	5 – 15	.418
Referential Literal accuracy	6.78 (1.35)	4 – 8	6.47 (1.27)	4 – 8	.382
Referential Metaphor accuracy	5.52 (1.76)	2 – 8	5.84 (1.78)	1 – 8	.509
PMM Metaphor accuracy	13.66 (3.49)	8 – 22	13.59 (5.40)	3 – 23	.961
Specificity of mental interpretation	4.78 (3.00)	-1 – 10	4.59 (2.92)	-2 – 10	.816
Post-training assessment					
Mental Strange Stories	8.52 (3.07)	3 – 14	8.81 (2.64)	4 – 14	–
Animated triangles	11.83 (2.33)	6 – 15	11.75 (2.54)	6 – 15	–
Referential Metaphor accuracy	6.26 (1.60)	4 – 8	6.72 (1.44)	2 – 8	–
PMM Metaphor accuracy	14.48 (3.44)	6 – 22	16.25 (4.33)	3 – 23	–
Specificity of mental interpretation	4.39 (2.87)	1 – 12	5.00 (3.15)	-3 – 10	–

Results showed that, at the pre-training assessment, children in the two training conditions did not significantly differ in any of the control and the focus variable (Table 4.2).

Training effects on ToM

Two-way mixed ANOVAs, with time (Pre-training and Post-training) as within-subjects factor and group (ToM and MetaCom) as between-subjects factor, showed a significant main effect of time, $F(1,53) = 6.24$, $p = .016$, $\eta_p^2 = .11$, but no time by group interaction, $F(1,53) < .01$, $p = .974$, $\eta_p^2 < .01$, on ToM assessed via the Strange Stories task. Similarly, we found a significant main effect of time, $F(1,53) = 7.80$, $p = .007$, $\eta_p^2 = .13$, but no time by group interaction, $F(1,53) = .65$, $p = .425$, $\eta_p^2 = .01$, on ToM assessed via the Triangle task.

Training effects on Metaphor Understanding

Two-way mixed ANOVAs, with time (Pre-training and Post-training) as within-subjects factor and group (ToM and MetaCom) as between-subjects factor, showed a significant main effect of time, $F(1,53) = 18.36$, $p < .001$, $\eta_p^2 = .26$, and a time by group interaction, $F(1,53) = 5.07$, $p = .029$, $\eta_p^2 = .09$, on metaphor accuracy assessed via the PMM task. Pairwise comparisons showed an improvement on metaphor accuracy in the MetaCom, $p < .001$, $d = .85$, 95%CI [1.60, 3.71], but not in the ToM group, $p = .188$, $d = .30$, 95% CI [-.42, 2.07]. Differently, we found a significant main effect of time, $F(1,53) = 12.84$, $p = .001$, $\eta_p^2 = .20$, but no time by group interaction, $F(1,53) = .09$, $p = .764$, $\eta_p^2 < .01$, on metaphor accuracy assessed via the referential task. Finally, on specificity of

mental interpretation, we found neither a significant main effect of time, $F(1,53) < .01$, $p = .986$, $\eta_p^2 < .01$, nor a time by group interaction, $F(1,53) = .87$, $p = .355$, $\eta_p^2 = .02$.

Discussion

The main aim of the present study was to investigate the different effects of the two training programs on the non-trained variables.

Before discussing the main results, it is important to highlight our preliminary findings concerning the psychometric properties of the hard version of the PMM task. Given that children's PMM scores were not at ceiling and showed adequate variability, it seems that even the hard version of the PMM task is a sound measure for detecting individual differences in metaphor understanding during middle childhood.

Regarding our main results and focusing on the effect on ToM, children in the MetaCom group did not differ from children in the ToM group in their ToM improvement. Given that previous studies have already shown that the ToM training was effective in enhancing ToM skills compared to an active control training (Bianco & Lecce, 2016; Bianco et al., 2016; Lecce & Bianco, 2018; Lecce et al., 2014), the present results suggest that both training programs (ToM and MetaCom) are effective in enhancing ToM skills in middle-aged children. Two characteristics of the MetaCom training could explain its role in enhancing ToM:

(1) The MetaCom training targeted the ability to make inferences about the metaphorical meaning within a story context. Thus, training this inferential

skill could have extended the MetaCoM training effect to the ability to make context-sensitive inferences about the main character's mental states in order to make sense of what is happening in a specific situation, namely ToM. In line with this, a previous study found that ToM in middle childhood was related to reading comprehension that included inferential questions (Cantin et al., 2016).

(2) Within the MetaCom training three stories included metaphors that conveyed a mental meaning. For this reason, the MetaCom training also enhanced children's ability to interpret a metaphor, when appropriate, as conveying mental features. Thus, training children's ability to prefer a mental interpretation of metaphors when possible and appropriate and to reason about mental states could, in turn, enhance ToM skills. This finding is supported by evidence suggesting that children's spontaneous tendency to prefer a mental, rather than a physical, interpretation of metaphor predicts later children's ToM ability (see Chapter 3). Accordingly, it is in line with the idea that being encouraged to refer to mental states promotes ToM throughout development (Bianco et al., 2016; Lecce et al., 2014).

Focusing on the training effect on metaphor accuracy, we found that only children in the MetaCom group improve in the comprehension of minimal-context nominal metaphors. On the contrary, children in the MetaCom group did not differ from children in the ToM group in their improvement in the comprehension of referential metaphors within a context. Given that previous studies have already shown that the MetaCom training was effective in enhancing metaphor understanding compared to an active control training (Tonini et al., 2020), the present results suggest that whereas only MetaCom training is effective in enhancing the understanding of minimal-context nominal metaphors, both training

(ToM and MetaCom) are effective in enhancing the comprehension of referential metaphors within a story-context. These results are in line with previous findings showing that the predictive effect of ToM on metaphor understanding was mainly driven by a broader inferential ability (see Chapter 3). Thus, training children's ability to infer the character's mental states within several different social scenarios would enhance children's broader ability to make inferences, also about metaphorical meaning, within a context. This hypothesis would explain why training children's ToM would enhance only children's ability to understand referential metaphors within a story-context and not children's ability to comprehend minimal-context metaphors. In line with this, previous studies found that ToM not only was related, but also affected children's later emerging reading comprehension, that required inferential comprehension skills (Atkinson, Slade, Powell, & Levy, 2017; Cantin et al., 2016).

Focusing on specificity of mental interpretation, children neither in the MetaCom group nor in the ToM group improved in specificity of mental interpretation after the training. The absence of a MetaCom training effect on specificity of mental interpretation was surprising but could be explained by the fact that only three stories within the MetaCom training have a mental content. In addition, the focus of the MetaCom training was not on training the ability to explicitly reason about the mental content in mental but not physical metaphors. Thus, a more focused and explicit training would better enhance the tendency to refer to mental aspects of the metaphor topic when possible and relevant (Durlak, Weissberg, Dymnicki, Taylor, & Schellinger, 2011). The absence of a ToM training effect on specificity of mental interpretation fits with previous evidence

suggesting that ToM did not predict later tendency to mentally interpret mental, but not physical metaphors (see Chapter 3).

Caveat and Conclusions

The present study contributes to clarifying the causal relationships between ToM and metaphor understanding, showing that whereas children in the ToM group improve in their ability to understand metaphors only when embedded in a context, children in the MetaCoM group improve in their ToM skills either when assessed children's ability to infer others' mental states or when measured children's tendency to attribute mental states.

However, the present study is not exempt from limitations. First of all, larger samples are needed to confirm our results. Second, we assessed children's improvements only immediately after the training. Further research should verify if these effects are long-lasting. Third, future studies should compare the effects of these two training (ToM and MetaCom) with different active control training programs. Indeed, comparing the effects of ToM and MetaCom training with the effect of a training targeting reading comprehension would help in understanding if the training effects on metaphor understanding and ToM are driven by an implicit effect of the training programs in understanding texts and make inferences about implicit information within stories. In addition, comparing the effect of ToM and MetaCoM training to an active control training targeting lexical and semantic knowledge would help understand the specificity of the present training effects.

Despite these limitations, the present study significantly contributes to our understanding of how ToM and pragmatics are related to each other and opens the way for future study to expand the scope of investigation considering additional pragmatics and ToM measures.

Besides the theoretical implications, the present study might also have the practical implication of promoting two fragile aspects of children's life, namely communicative and socio-cognitive skills. For instance, on the one side training socio-cognitive skills would help children, not only to understand others' mental states, but also to make inferences from implicit information within a story and, on the other side training metaphor understanding could help children, not only to make pragmatic inferences but also put themselves into someone else's shoes.

General discussion and conclusions

Metaphor understanding is an advanced pragmatic skill that requires the ability to fill the gap between the literal and the intended meaning of a social partner (Carston, 2012; Sperber & Wilson, 2002). Usually, communicative exchanges involving metaphors require the activation of shared social knowledge and higher sensitivity to social cues (Bowes & Katz, 2015). Metaphors are therefore important during social exchanges since they create intimacy between the listener and the speakers by drawing them closer to one another (Bowes & Katz, 2015; Cohen, 1978; Horton, 2007). Only literature on adulthood investigated the social function of metaphors. The lack of studies on middle childhood is striking, given that middle childhood is a key period for both social relationships and metaphor understanding (Noveck et al., 2001; Rubin et al., 2007; Winner et al., 1976). Thus, the first study of the present dissertation aimed to fill this gap by exploring the existence and the direction of the developmental associations between metaphor understanding and social relationships focusing on the crucial developmental stage of middle childhood.

Results of Study 1 showed a longitudinal and bidirectional association between metaphor understanding and peer rejection (Table 5.1). On the one hand, previous studies suggested that peer rejection, being a stressful life event, depleted individuals' cognitive inner resources and limited children's development of cognitive functions (Lecce et al., 2020; Muraven & Baumeister, 2000; Muraven et al., 1998). Starting from this claim we, therefore, argue that peer rejection hinders the development of several cognitive abilities including metaphor understanding. On the other hand, recent empirical evidence on pragmatic conversational skills in

pre-school children suggested that having poor pragmatic skills led to be less attractive playmates and, consequently, to become highly rejected (van der Wilt et al., 2018, 2020). Thus, our findings extend these results to middle childhood, and to a higher inferential pragmatic skill, namely metaphor understanding. We conclude that understanding the intended meaning in non-literal communicative exchanges (as in metaphors) is an ability that leads to being less attractive playmates since this skill is not only important but necessary for social relationships.

Finally, we found that the longitudinal associations between peer rejection and metaphor understanding had an equal strength in the two directions, suggesting that the role of being rejected in hindering the development of metaphor comprehension and the one of poor metaphorical skills in increasing the risk of being rejected by peers are equally stronger.

Starting from the evidence on the link between peer rejection and metaphor understanding, and given the literature showing a relationship between peer rejection and ToM (Banerjee et al., 2011), the second study aimed to explore the association between ToM and metaphor understanding. The link between ToM and metaphor understanding was little explored, especially in typical development. In addition, the few studies that explored this issue did not distinguish between mental and physical metaphors, which seems to be crucial to unpack the broader association between ToM and pragmatics (Matthews et al., 2018). Results showed that 9-year-olds, but not older children, who were better in ToM were also more likely to better interpret mental, but not physical, metaphors (Table 5.1). Study 2 has two main implications. First, it suggests that the

association between metaphor interpretation and ToM is not general, but rather specific for those metaphors whose interpretation requires an inference about mental states, supporting the theoretical distinction between psychological and physical metaphors (Melogno et al., 2017; Vosniadou et al., 1984; Winner et al., 1976). Second, it suggests that the relationship between metaphor interpretation and ToM is stronger in earlier developmental phases. This fits with the literature about the relationship between ToM and language that was found to be stronger in pre-schoolers and seemed to become weaker across middle childhood (Lecce et al., 2010; Milligan et al., 200).

Although interesting and pioneering, the cross-sectional nature of these results prevents us to make conclusions about the longitudinal associations between ToM and metaphor understanding. For this reason, the third study tried to overcome this limit by adopting a cross-lagged longitudinal design to examine not simply the existence but also the direction of the associations between individual differences in ToM and metaphor understanding.

The first result of Study 3 showed that in typically developing middle-aged children metaphor comprehension was bi-directionally linked to general inferential skill (i.e., ability to make inferences about physical states), but not ToM (i.e., ability to make inferences about mental states) (Table 5.1). Our findings fit with evidence coming from the literature on adult clinical populations finding an association between pragmatics and non-mental-specific inferential skills (Martin & McDonald, 2005; Montemurro et al., 2019). Overall, this pattern of results suggests that the ability to make inferences, not specifically about mental states, is strongly implicated in the development of metaphor

comprehension and vice versa. In addition, we claim that the role of metaphor accuracy in ToM development could be, at least in part, explained by a general inferential ability.

The second result of the third study showed that the tendency to interpret mentally mental, but not physical, metaphors (i.e., specificity of mental interpretation) predicted later ToM, but not vice versa (Table 5.1). In line with this, a robust number of studies showed that mental state lexicon predicted later ToM but not vice versa (Hughes & Dunn, 1998; Ruffman et al., 2002). Our findings add to this literature that, not only talking about mental states but also talking about mental metaphors might drive ToM development.

Overall, Study 3 of the present dissertation expands the results of study 2, adding a developmental angle to the analysis of the associations between ToM and metaphors. This third study supported the hypothesis according to which the direction of the developmental associations between ToM and metaphor understanding varies depending on the pragmatic aspect considered, namely accuracy vs. specificity of mental interpretation, and is in part explained by the role of general inferential ability.

Since the results of Study 3, coming from a longitudinal design, were able to demonstrate genuine developmental relationships but not causal relationships (Bradley & Bryant, 1983), the fourth study tried to answer this last question by adopting a training design. Thus, the aim of Study 4 was to investigate the causality of the relationships between ToM and metaphor understanding in 9-year-old typically developing children.

The first result of Study 4 showed that a training about metaphor understanding, that targets both metaphor comprehension and metaphor interpretation, generalized its effect of ToM (Table 5.1). This result fits with findings coming from study three. On the one hand, the ability to make inferences about metaphorical meaning not only predicts but also prompts the development of general inferential skills, including inference about mental states, namely ToM. On the other hand, interpret a metaphor, when appropriate, as conveying mental features not only predicts later ToM but has also a causal role in enhancing ToM. In line with this view, studies on middle childhood reported that being prone to refer to mental states promotes ToM throughout development (Bianco et al., 2016; Lecce, et al. 2014).

The second result of Study 4 showed that a ToM training generalized its effect on the ability to understand referential metaphor within a context but not on the ability to understand minimal-context nominal metaphor (Table 5.1). This result expanded the findings of Study 3, showing that the predictive role of ToM on the development of metaphor comprehension was mainly driven by the general ability to make inferences. Thus, training children's ability to infer the character's mental states within several different social scenarios would enhance children's broader ability to make inferences, also about metaphorical meaning, within a story-context. However, since the ability to make inferences about mental states was trained within a story context, the training did not generalize its effect on the ability to make inferences about the most salient metaphorical meaning of a sentence. For an overview of all the studies presented in the present dissertation see Table 5.1.

From a more general point of view, the results coming from Studies 2, 3, and 4 endorse the paper arguing that generally speaking, pragmatics and ToM do not overlap (Bosco et al., 2018). Indeed, echoing the theoretical models discussed in Study 3, the relationships between ToM and pragmatics, specifically metaphor understanding, are complex and vary depending on several different factors, including the pragmatic phenomena considered (Andrés-Roqueta & Katsos, 2017; Matthews et al., 2018). Thus, Studies 2, 3, and 4 suggest that while the interpretation of mental, but not physical, metaphors is linked with ToM, the link between ToM and metaphor comprehension is mainly driven by the general ability to make inferences that underlies both skills. Focusing on the role of ToM in the development of the broader pragmatic ability, it seems that the contribution of ToM is due to the ability to make inferences in general and not only to the ability to make specific inferences about mental states. Focusing on the role of pragmatics in the development of ToM, the studies described in the present dissertation suggest that pragmatic inferences drive the development of ToM use in social situations, e.g. during conversations. In line with this claim, conversations are a thriving environment for developing ToM. For example, communicative exchanges involving metaphors require the listener to activate shared social knowledge and increase his/her sensitivity to the speaker's minds (Bowes & Katz, 2015). Thus, the present discussion fits with the claim according to which conversations allow to reflect on the idea that people could have different social experiences and different mental states about the same situation and, during development, make children enter the 'community of minds' (Nelson, 2007).

Table 5.1. Main aim, design, sample, and main results of the studies reported in the present dissertation.

Study	Main aim	Design	Sample	Results
Study1	To investigate the *existence* and the *direction* of the associations between metaphor understanding and social relationships	1-year longitudinal	126 typically developing children from 9 to 10 years	Peer rejection, but not peer acceptance, was longitudinally and bi-directionally related to metaphor understanding
Study2	To investigate the *existence* of the relationship between individual differences in ToM and metaphor understanding	Cross-sectional	217 typically developing children: - 62 9-year-olds - 48 10-year-olds - 51 11-year-olds - 56 12-year-olds	ToM and interpretation of mental metaphors were associated. This link was stronger for earlier phases of development
Study3	To investigate the *direction* of the relationships between individual differences in ToM and metaphor understanding	6-months longitudinal	54 typically developing 9-year-old children	- Specificity of mental interpretation of metaphors predicted later ToM but not vice versa - Metaphor Accuracy was bi-directionally linked with general inferential skill
Study 4	To investigate the *causal* relationships between individual differences in ToM and metaphor understanding	Training	55 typically developing 9-year-old children - 23 attended the ToM training - 32 attended the MetaCom training	- MetaCom training extended its effect on ToM - ToM training extended its effect on the understanding of referential metaphors within a context but not on the understanding of minimal-context nominal metaphors

Future directions

The present dissertation reported four pioneering studies that shed light on the associations between metaphor understanding and both social relationships and ToM in middle childhood. However future studies are needed to deepen the knowledge in this research topic.

First of all, concerning the methodology, longitudinal and training studies that examine at the same time the associations between metaphor understanding, social relationships, and ToM in middle childhood are needed. This would help to clarify the mutual influences between these three related constructs.

Second, concerning the measures, future studies should consider other metaphorical skills that may be differently related to ToM, such as metaphor production. In this respect, a recent position paper claimed that pretend play and perspective-taking, which are known to be ToM precursors, are forerunners of similes and metaphor production as they reflect the ability to consider more than one mental representation at the time (Clark, 2019). Interestingly, according to this view, metaphor production should be predicted by ToM skills and not vice versa. In addition, the present study could prompt future researchers to expand our scope by examining the relationships between ToM and pragmatics in children, by looking at other pragmatic skills, that might differently capitalize on ToM. For instance, the present dissertation could prompt future researchers to deepen the knowledge about the relationships between ToM and irony and humor (e.g., Bischetti, Ceccato, Lecce, Cavallini, & Bambini, 2019; Bosco & Gabbatore, 2017; Massaro, Valle, & Marchetti, 2014). Here, a very recent study on elderly people

showed that general pragmatic skills were always involved in understanding jokes, but ToM skills were involved only when reasoning about the joke characters' mental states was required (Bischetti et al., 2019).

Finally, regarding the developmental stage under examination, future studies should investigate how the relationships found in middle childhood occur in the same manner in other developmental phases, such as in aging and pre-school years.

References

Abkarian, G. G., Jones, A., & West, G. (1992). Young Children's Idiom Comprehension. *Journal of Speech, Language, and Hearing Research, 35*(3), 580-587. doi:doi:10.1044/jshr.3503.580

Adamczyk, P., Daren, A., Sułecka, A., Błądziński, P., Cichocki, L., Kalisz, A., Gawęda, L., & Cechnicki, A. (2016). Do better communication skills promote sheltered employment in schizophrenia?. *Schizophrenia Research*, 176(2–3), 331-339. doi: 10.1016/j.schres.2016.08.015

Angeleri, R., & Airenti, G. (2014). The development of joke and irony understanding: a study with 3- to 6-year-old children. *Canadian Journal of Experimental Psychology, 68*(2), 133-146. doi:10.1037/cep0000011

Appleton, M., & Reddy, V. (1996). Teaching Three Year-Olds to Pass False Belief Tests: A Conversational Approach. *Social Development, 5*(3), 275-291. doi:10.1111/j.1467-9507.1996.tb00086.x

Arcara, G., & Bambini, V. (2016). A Test for the Assessment of Pragmatic Abilities and Cognitive Substrates (APACS): Normative Data and Psychometric Properties. *Frontiers in Psychology, 7*, 70. doi:10.3389/fpsyg.2016.00070

American Psychiatric Association (2013). *Diagnostic and statistical manual of mental disorders (DSM-5®)*: American Psychiatric Pub.

Atkinson, L., Slade, L., Powell, D., & Levy, J. P. (2017). Theory of mind in emerging reading comprehension: A longitudinal study of early indirect

and direct effects. *Journal of Experimental Child Psychology, 164*, 225-238. doi:10.1016/j.jecp.2017.04.007

Austin, G., Groppe, K., & Elsner, B. (2014). The reciprocal relationship between executive function and theory of mind in middle childhood: a 1-year longitudinal perspective. *Frontiers in Psychology, 5*(655). doi:10.3389/fpsyg.2014.00655

Babad, E. (2001). On the conception and measurement of popularity: more facts and some straight conclusions. *Social Psychology of Education, 5*(1), 3-29. doi: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. doi:10.1111/sode.12056

Baer, C., & Friedman, O. (2018). Fitting the Message to the Listener: Children Selectively Mention General and Specific Facts. *Child Development, 89*(2), 461-475. doi:10.1111/cdev.12751

Bagwell, C. L., Coie, J. D., Terry, R. A., & Lochman, J. E. (2000). Peer clique participation and social status in preadolescence. *Merrill-Palmer Quarterly, 46*(2), 280-305.

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. doi: 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. doi:10.1111/j.1467-8624.2011.01669.x

Barner, D., Brooks, N., & Bale, A. (2011). Accessing the unsaid: The role of scalar alternatives in children's pragmatic inference. *Cognition, 118*(1), 84-93. doi:10.1016/j.cognition.2010.10.010

Baron-Cohen, S. (2000). Theory of mind and autism: A review *International Review of Research in Mental Retardation* (Vol. 23, pp. 169-184): Academic Press.

Baron-Cohen, S., Leslie, A. M., & Frith, U. (1985). Does the autistic child have a "theory of mind" ? *Cognition, 21*(1), 37-46. doi:10.1016/0010-0277(85)90022-8

Berenguer, C., Miranda, A., Colomer, C., Baixauli, I., & Roselló, B. (2018). Contribution of Theory of Mind, Executive Functioning, and Pragmatics to Socialization Behaviors of Children with High-Functioning Autism. *Journal of autism and Developmental disorders, 48*(2), 430-441. doi:10.1007/s10803-017-3349-0

Bernicot, J., Laval, V., & Chaminaud, S. (2007). Nonliteral language forms in children: In what order are they acquired in pragmatics and metapragmatics? *Journal of Pragmatics, 39*(12), 2115-2132. doi:10.1016/j.pragma.2007.05.009

Bernstein, D. M., Atance, C., Loftus, G. R., & Meltzoff, A. (2004). We saw it all along: visual hindsight bias in children and adults. *Psychological science, 15*(4), 264-267. doi:10.1111/j.0963-7214.2004.00663.x

Białecka-Pikul, M. (2010). Teaching children to understand metaphors as a path leading to theory of mind development. *European Journal of Developmental Psychology, 7*(5), 529-544. doi:10.1080/17405620802643692

Bianco, F., & Lecce, S. (2016). Translating child development research into practice: Can teachers foster children's theory of mind in primary school? *British Journal of Educational Psychology, 86*(4), 592-605. doi:10.1111/bjep.12125

Bianco, F., Lecce, S., & Banerjee, R. (2016). Conversations about mental states and theory of mind development during middle childhood: A training study. *Journal of Experimental Child Psychology, 149*, 41-61. doi:10.1016/j.jecp.2015.11.006

Bierman, K. L. (2004). *Peer rejection: Developmental processes and intervention strategies.* New York, US: Guilford Press.

Bierman, K. L., & Furman, W. (1984). The Effects of Social Skills Training and Peer Involvement on the Social Adjustment of Preadolescents. *Child Development, 55*(1), 151. doi:10.2307/1129841

Bischetti, L., Ceccato, I., Lecce, S., Cavallini, E., & Bambini, V. (2019). Pragmatics and theory of mind in older adults' humor comprehension. *Current Psychology*. doi:10.1007/s12144-019-00295-w

Bishop, D. (1998). Development of the Children's Communication Checklist (CCC): a method for assessing qualitative aspects of communicative impairment in children. *Journal of Child Psychology and Psychiatry, 39*(6), 879-891.

Bishop, D. (2003). *Test for Reception of Grammar version 2 (TROG-2)*: Giunti Psychometrics.

Bishop, D. (2013). *Children's Communication Checklist (CCC-2)*. In F. R. Volkmar (Ed.), Encyclopedia of Autism Spectrum Disorders (pp. 614-618). New York, NY: Springer New York.

Black, B., & Hazen, N. L. (1990). Social status and patterns of communication in acquainted and unacquainted preschool children. *Developmental Psychology, 26*(3), 379-387. doi:10.1037/0012-1649.26.3.379

Black, B., & Logan, A. (1995). Links between Communication Patterns in Mother-Child, Father-Child, and Child-Peer Interactions and Children's Social Status. *Child Development, 66*(1), 255-271. doi:10.1111/j.1467-8624.1995.tb00869.x

Blain-Brière, B., Bouchard, C., & Bigras, N. (2014). The role of executive functions in the pragmatic skills of children age 4–5. *Frontiers in Psychology, 5*(240). doi:10.3389/fpsyg.2014.00240

Bohn, M., & Frank, M. C. (2019). The Pervasive Role of Pragmatics in Early Language. *Annual Review of Developmental Psychology, 1*(1), 223-249. doi:10.1146/annurev-devpsych-121318-085037

Bosco, F. M., & Gabbatore, I. (2017). Sincere, Deceitful, and Ironic Communicative Acts and the Role of the Theory of Mind in Childhood. *Frontiers in Psychology, 8*(21). doi:10.3389/fpsyg.2017.00021

Bosco, F. M., Tirassa, M., & Gabbatore, I. (2018). Why Pragmatics and Theory of Mind Do Not (Completely) Overlap. *Frontiers in Psychology, 9*, 1453. doi:10.3389/fpsyg.2018.01453

Bowes, A., & Katz, A. (2015). Metaphor creates intimacy and temporarily enhances theory of mind. *Memory & Cognition, 43*(6), 953–963. doi:10.3758/s13421-015-0508-4

Bradley, L., & Bryant, P. E. (1983). Categorizing sounds and learning to read—a causal connection. *Nature, 301*(5899), 419-421. doi:10.1038/301419a0

Brown, J. R., Donelan-McCall, N., & Dunn, J. (1996). Why Talk about Mental States? The Significance of Children's Conversations with Friends, Siblings, and Mothers. *Child Development, 67*(3), 836-849. doi:10.2307/1131864

Brown, T. A. (2015). *Confirmatory Factor Analysis for Applied Research, Second Edition*: Guilford Publications.

Buhs, E. S., & Ladd, G. W. (2001). Peer rejection as antecedent of young children's school adjustment: An examination of mediating processes. *Developmental Psychology, 37*(4), 550–560. doi:10.1037/0012-1649.37.4.550

Burleson, B. R., Delia, J. G., & Applegate, J. L. (1992). Effects of Maternal Communication and Children's Social-Cognitive and Communication

Skills on Children's Acceptance by the Peer Group. *Family Relations, 41*(3), 264-272. doi:10.2307/585189

Caillies, S., & Le Sourn-Bissaoui, S. (2008). Children's understanding of idioms and theory of mind development. *Developmental Science, 11*(5), 703-711. doi:10.1111/j.1467-7687.2008.00720.x

Cameron, L. (2008). Metaphor and talk. In J. Gibbs & W. Raymond (Eds.), *The Cambridge Handbook of Metaphor and Thought* (pp. 197-211). Cambridge: Cambridge University Press.

Cantin, R. H., Gnaedinger, E. K., Gallaway, K. C., Hesson-McInnis, M. S., & Hund, A. M. (2016). Executive functioning predicts reading, mathematics, and theory of mind during the elementary years. *Journal of Experimental Child Psychology, 146*, 66-78. doi:10.1016/j.jecp.2016.01.014

Caputi, M., Lecce, S., Pagnin, A., & Banerjee, R. (2012). Longitudinal effects of theory of mind on later peer relations: The role of prosocial behavior. *Developmental Psychology*, 48(1), 257-270. doi:10.1037/a0025402

Carpendale, J. I., & Chandler, M. J. (1996). On the distinction between false belief understanding and subscribing to an interpretive theory of mind. *Child Development, 67*(4), 1686-1706. doi:10.2307/1131725

Carriedo, N., Corral, A., Montoro, P. R., Herrero, L., Ballestrino, P., & Sebastián, I. (2016). The development of metaphor comprehension and its relationship with relational verbal reasoning and executive function. *PLoS ONE, 11*(3). doi:10.1371/journal.pone.0150289

Carston, R. (2012). Metaphor and the literal/non-literal distinction. In K. M. Jaszczolt & K. Allan (Eds.), *The Cambridge Handbook of Pragmatics* (pp. 469-492). Cambridge: Cambridge University Press.

Castelli, F., Happé, F., Frith, U., & Frith, C. (2000). Movement and Mind: A Functional Imaging Study of Perception and Interpretation of Complex Intentional Movement Patterns. *NeuroImage, 12*(3), 314-325. doi:10.1006/nimg.2000.0612

Champagne-Lavau, M., & Stip, E. (2010). Pragmatic and executive dysfunction in schizophrenia. *Journal of Neurolinguistics, 23*(3), 285-296. doi:10.1016/j.jneuroling.2009.08.009

Chen, Q., Hughes, J. N., Liew, J., & Kwok, O.-M. (2010). Joint contributions of peer acceptance and peer academic reputation to achievement in academically at-risk children: Mediating processes. Journal of Applied *Developmental Psychology, 31*(6), 448–459. doi:10.1016/j.appdev.2010.09.001

Cheung, H. S., & Elliott, J. M. (2017). Child shyness and peer likeability: The moderating role of pragmatics and vocabulary. *British Journal of Developmental Psychology, 35*(4), 531-545. doi:10.1111/bjdp.12192

Cillessen, A. H. N., & Marks, P. E. L. (2017). Methodological Choices in Peer Nomination Research. *New Directions for Child and Adolescent Development, 2017*(157), 21-44. doi:10.1002/cad.20206

Clark, E. V. (2019). Perspective-taking and pretend-play: Precursors to figurative language use in young children. *Journal of Pragmatics*. doi:10.1016/j.pragma.2018.12.012

Clements, W. A., Rustin, C. L., & McCallum, S. (2000). Promoting the Transition from Implicit to Explicit Understanding: A Training Study of False Belief. *Developmental Science, 3*(1), 81-92. doi:10.1111/1467-7687.00102

Cohen, T. (1978). Metaphor and the Cultivation of Intimacy. *Critical Inquiry, 5*(1), 3-12.

Coie, J. D. (1990). Toward a theory of peer rejection. In S.R. Asher & J.D. Coie (Eds.), *Peer rejection in childhood* (pp. 365-401). Cambridge, England: Cambridge University Press.

Coie, J. D., Dodge, K. A., & Coppotelli, H. (1982). Dimensions and types of social status: A cross-age perspective. *Developmental Psychology, 18*(4), 557-570. doi:10.1037/0012-1649.18.4.557

Conti-Ramsden, G., & Botting, N. (2004). Social Difficulties and Victimization in Children With SLI at 11 Years of Age. *Journal of Speech, Language, and Hearing Research, 47*(1), 145-161. doi:doi:10.1044/1092-4388(2004/013)

Cornoldi, C., & Colpo, G. (1998). *Prove di lettura MT per la scuola elementare 2 [MT Tests of Reading for Primary School 2]*. Florence, Italy: Organizzazioni Speciali.

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. doi:10.1016/j.socscimed.2007.11.024

Değirmencioğlu, S. M., Urberg, K. A., Tolson, J. M., & Richard, P. (1998). Adolescent friendship networks: Continuity and change over the school year. *Merrill-Palmer Quarterly, 44*(3), 313-337.

Demorest, A., Meyer, C., Phelps, E., Gardner, H., & Winner, E. (1984). Words Speak Louder Than Actions: Understanding Deliberately False Remarks. *Child Development, 55*(4), 1527-1534. doi:10.2307/1130022

Demorest, A., Silberstein, L., Gardner, H., & Winner, E. (1983). Telling it as it isn't: Children's understanding of figurative language. *British Journal of Developmental Psychology, 1*(2), 121-134. doi:10.1111/j.2044-835X.1983.tb00550.x

Dent, C., & Rosenberg, L. (1990). Visual and verbal metaphors: developmental interactions. *Child Development, 61*(4), 983-994. doi:10.1111/j.1467-8624.1990.tb02836.x

Devine, R. T., & Hughes, C. (2013). Silent films and strange stories: theory of mind, gender, and social experiences in middle childhood. *Child Development, 84*(3), 989-1003. doi:10.1111/cdev.12017

Devine, R. T., White, N., Ensor, R., & Hughes, C. (2016). Theory of mind in middle childhood: Longitudinal associations with executive function and social competence. *Developmental Psychology, 52*(5), 758-771. doi:10.1037/dev0000105

Dews, S., & Winner, E. (1997). Chapter 10 Attributing meaning to deliberately false utterances: The case of irony. In C. Mandell & A. McCabe (Eds.), *Advances in Psychology* (Vol. 122, pp. 377-414): North-Holland.

Dumontheil, I., Apperly, I. A., & Blakemore, S. J. (2010). Online usage of theory of mind continues to develop in late adolescence. *Developmental Science, 13*(2), 331-338. doi:10.1111/j.1467-7687.2009.00888.x

Dunn, J. (1999). Making sense of the social world: Mindreading, emotion, and relationships *Developing theories of intention: Social understanding and self-control.* (pp. 229-242). Mahwah, NJ, US: Lawrence Erlbaum Associates Publishers.

Dunn, L. M., & Dunn, L. M. (1981). *Peabody Picture Vocabulary Test-Revised (PPVT-R).* . Circle Pines, Minnesota: American Guidance Sevice.

Durlak, J. A., Weissberg, R. P., Dymnicki, A. B., Taylor, R. D., & Schellinger, K. B. (2011). The Impact of Enhancing Students' Social and Emotional Learning: A Meta-Analysis of School-Based Universal Interventions. *Child Development, 82*(1), 405-432. doi:10.1111/j.1467-8624.2010.01564.x

Ebert, S. (2020). Theory of mind, language, and reading: Developmental relations from early childhood to early adolescence. *Journal of Experimental Child Psychology, 191*, 104739. doi:10.1016/j.jecp.2019.104739

Eckerman, C. O. (1979). The human infant in social interaction. In R. Cairns (Ed.), *The analysis of social interactions: Methods, issues, and illustrations* (pp. 163–178). Hillsdale, NJ: Erlbaum.

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. doi:10.1207/S15328007SEM0803_5

Ensor, R., & Hughes, C. (2008). Content or connectedness? Mother-child talk and early social understanding. *Child Development, 79*(1), 201-216. doi:10.1111/j.1467-8624.2007.01120.x

Falkum, I. (2019). Pragmatic development: Learning to use language to communicate. In J. Horst & J. von Koss Torkildsen (Eds.), *International handbook of language acquisition.* Routledge, Abingdon. doi:10.4324/9781315110622-13

Filippova, E., & Astington, J. W. (2008). Further development in social reasoning revealed in discourse irony understanding. *Child Development, 79*(1), 126-138. doi:10.1111/j.1467-8624.2007.01115.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. doi: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. doi:10.1007/s40894-016-0042-z

Frank, M. C., & Goodman, N. D. (2014). Inferring word meanings by assuming that speakers are informative. *Cognitive Psychology, 75*, 80-96. doi:10.1016/j.cogpsych.2014.08.002

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. doi:10.1038/srep40871

Gifford-Smith, M. E., & Brownell, C. A. (2003). Childhood peer relationships: social acceptance, friendships, and peer networks. *Journal of School Psychology, 41*(4), 235-284. doi:10.1016/S0022-4405(03)00048-7

Giora, R. (2003). *On our mind: Salience, context, and figurative language*. New York, NY, US: Oxford University Press.

Gleitman, L. (1990). The Structural Sources of Verb Meanings. *Language Acquisition, 1*(1), 3-55.

Glucksberg, S. (1989). Metaphors in Conversation: How Are They Understood? Why Are They Used? *Metaphor and Symbolic Activity, 4*(3), 125-143. doi:10.1207/s15327868ms0403_2

Göncü, A. (1993). Development of intersubjectivity in the dyadic play of preschoolers. *Early Childhood Research Quarterly, 8*(1), 99-116. doi:10.1016/S0885-2006(05)80100-0

Goodman, R. (1997). The Strengths and Difficulties Questionnaire: a research note. *Journal of Child Psychology and Psychiatry, 38*(5), 581-586.

Graham, S. A., Sedivy, J., & Khu, M. (2013). That's not what you said earlier: preschoolers expect partners to be referentially consistent. *Journal of Child Language, 41*(1), 34-50. doi:10.1017/S0305000912000530

Gredebäck, G., Astor, K., & Fawcett, C. (2018). Gaze Following Is Not Dependent on Ostensive Cues: A Critical Test of Natural Pedagogy. *Child Development, 89*(6), 2091-2098. doi:10.1111/cdev.13026

Halberda, J. (2003). The development of a word-learning strategy. *Cognition, 87*(1), B23-B34. doi:10.1016/S0010-0277(02)00186-5

Happé, F. G. (1993). Communicative competence and theory of mind in autism: a test of relevance theory. *Cognition, 48*(2), 101-119. doi:10.1016/0010-0277(93)90026-R

Happé, F. G. (1994). An advanced test of theory of mind: Understanding of story characters' thoughts and feelings by able autistic, mentally handicapped, and normal children and adults. *Journal of autism and Developmental disorders, 24*(2), 129-154. doi:10.1007/BF02172093

Harris, M., & Pexman, P. M. (2003). Children's perceptions of the social functions of verbal irony. *Discourse Processes, 36*(3), 147-165. doi:10.1207/S15326950DP3603_1

Hawley, P. H., Little, T. D., & Pasupathi, M. (2002). Winning friends and influencing peers: Strategies of peer influence in late childhood. *International Journal of Behavioral Development, 26*(5), 466-474. doi:10.1080/01650250143000427

Hay, D. F., Nash, A., & Pedersen, J. (1983). Interaction between six-month-old peers. *Child Development, 54*(3), 557-562. doi:10.2307/1130042

Hay, D. F., Payne, A., & Chadwick, A. (2004). Peer relations in childhood. *Journal of Child Psychology and Psychiatry, 45*(1), 84-108. doi:10.1046/j.0021-9630.2003.00308.x

Hazen, N. L., & Black, B. (1989). Preschool Peer Communication Skills: The Role of Social Status and Interaction Context. *Child Development, 60*(4), 867-876. doi:10.2307/1131028

Hernik, M., & Broesch, T. (2019). Infant gaze following depends on communicative signals: An eye-tracking study of 5- to 7-month-olds in Vanuatu. *Developmental Science, 22*(4), e12779. doi:10.1111/desc.12779

Hinshaw, S. P. (1992). Externalizing behavior problems and academic underachievement in childhood and adolescence: Causal relationships and underlying mechanisms. *Psychological Bulletin, 111*(1), 127-155. doi:10.1037/0033-2909.111.1.127

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

Howes, C. (1988). Peer interaction of young children. *Monographs of the Society for Research in Child Development, 53*(1), 94-94. doi:10.2307/1166062

Howes, C., Unger, O. A., & Matheson, C. C. (1992). *The collaborative construction of pretend: Social pretend play functions*. Albany, NY, US: State University of New York Press.

Huang, S., Oi, M., & Taguchi, A. (2015). Comprehension of figurative language in Taiwanese children with autism: The role of theory of mind and receptive vocabulary. *Clinical Linguistics & Phonetics, 29*(8-10), 764-775. doi:10.3109/02699206.2015.1027833

Huang, Y. T., & Snedeker, J. (2009). Online interpretation of scalar quantifiers: Insight into the semantics–pragmatics interface. *Cognitive Psychology, 58*(3), 376-415. doi:10.1016/j.cogpsych.2008.09.001

Hübscher, I., & 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. (1998). Executive function in preschoolers: Links with theory of mind and verbal ability. *British Journal of Developmental Psychology, 16*(2), 233-253. doi:10.1111/j.2044-835X.1998.tb00921.x

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. doi:10.1016/j.jecp.2016.01.017

Hughes, C., & Devine, R. T. (2015). Individual differences in theory of mind from preschool to adolescence: Achievements and directions. *Child Development Perspectives, 9*(3), 149-153. doi:10.1111/cdep.12124

Hughes, C., & Dunn, J. (1998). Understanding mind and emotion: longitudinal associations with mental-state talk between young friends. *Developmental Psychology, 34*(5), 1026-1037. doi:10.1037//0012-1649.34.5.1026

Hughes, C., Lecce, S., & Wilson, C. (2007). "Do you know what I want?" Preschoolers' talk about desires, thoughts and feelings in their conversations with sibs and friends. *Cognition and Emotion, 21*(2), 330-350. doi:10.1080/02699930600551691

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

Kalandadze, T., Bambini, V., & Næss, K. B. (2019). A systematic review and meta-analysis of studies on metaphor comprehension in individuals with autism spectrum disorder: Do task properties matter? *Applied Psycholinguistics, 40*(6), 1421-1454. doi:10.1017/S0142716419000328

Kalandadze, T., Norbury, C., Nærland, T., & Næss, K. B. (2016). Figurative language comprehension in individuals with autism spectrum disorder: A meta-analytic review. *Autism, 22*(2), 99-117. doi:10.1177/1362361316668652

Karmiloff-Smith, A. (1995). *Beyond Modularity: A Developmental Perspective on Cognitive Science*: A Bradford Book.

Kline, R. B. (2011). *Principles and practice of structural equation modeling, 3rd ed*. New York, NY, US: Guilford Press.

Köymen, B., Mammen, M., & Tomasello, M. (2016). Preschoolers use common ground in their justificatory reasoning with peers. *Developmental Psychology, 52*(3), 423-429. doi:10.1037/dev0000089

Kraatz-Keiley, M., Bates, J. E., Dodge, K. A., & Pettit, G. S. (2000). A Cross-Domain Growth Analysis: Externalizing and Internalizing Behaviors During 8 Years of Childhood. *Journal of Abnormal Child Psychology, 28*(2), 161-179. doi:10.1023/A:1005122814723

Krasnor, L. R., & Rubin, K. H. (1983). Preschool Social Problem Solving: Attempts and Outcomes in Naturalistic Interaction. *Child Development, 54*(6), 1545-1558. doi:10.2307/1129818

Kuder, G. F., & Richardson, M. W. (1937). The theory of the estimation of test reliability. *Psychometrika, 2*(3), 151-160. doi:10.1007/BF02288391

Ladd, G. W. (1981). Effectiveness of a social learning method for enhancing children's social interaction and peer acceptance. *Child Development, 52*(1), 171-178. doi:10.2307/1129227

Ladd, G. W., & Troop-Gordon, W. (2003). The role of chronic peer difficulties in the development of children's psychological adjustment problems. *Child Development, 74*(5), 1344-1367. doi:10.1111/1467-8624.00611

Lagattuta, K. H., Elrod, N. M., & Kramer, H. J. (2016). How do thoughts, emotions, and decisions align? A new way to examine theory of mind during middle childhood and beyond. *Journal of Experimental Child Psychology, 149*, 116-133. doi:10.1016/j.jecp.2016.01.013

Lagattuta, K. H., Sayfan, L., & Blattman, A. J. (2010). Forgetting common ground: Six- to seven-year-olds have an overinterpretive theory of mind. *Developmental Psychology, 46*(6), 1417-1432. doi:10.1037/a0021062

Lagattuta, K. H., Sayfan, L., & Harvey, C. (2014). Beliefs About Thought Probability: Evidence for Persistent Errors in Mindreading and Links to Executive Control. *Child Development, 85*(2), 659-674. doi:10.1111/cdev.12154

Landau, M. J., Meier, B. P., & Keefer, L. A. (2010). A metaphor-enriched social cognition. *Psychological Bulletin, 136*, 1045–1067. doi:10.1037/a0020970

Larson, J. J., Whitton, S. W., Hauser, S. T., & Allen, J. P. (2007). Being close and being social: peer ratings of distinct aspects of young adult social competence. *Journal of Personality Assessment, 89*(2), 136-148. doi:10.1080/00223890701468501

Law, J., Rush, R., Clegg, J., Peters, T., & Roulstone, S. (2015). The role of pragmatics in mediating the relationship between social disadvantage and adolescent behavior. *Journal of Developmental & Behavioral Pediatrics, 36*(5), 389-398. doi:10.1097/dbp.0000000000000180

Laws, G., & Bishop, D. V. M. (2004). Pragmatic language impairment and social deficits in Williams syndrome: A comparison with Down's syndrome and specific language impairment. *International Journal of Language & Communication Disorders, 39*(1), 45-64. doi:10.1080/13682820310001615797

Lease, A. M., Kennedy, C. A., & Axelrod, J. L. (2002). Children's Social Constructions of Popularity. *Social Development, 11*(1), 87-109. doi:10.1111/1467-9507.00188

Lecce, S., & Bianco, F. (2018). Working memory predicts changes in children's theory of mind during middle childhood: A training study. *Cognitive Development, 47*, 71-81. doi:10.1016/j.cogdev.2018.04.002

Lecce, S., Bianco, F., Devine, R. T., & Hughes, C. (2017). Relations between theory of mind and executive function in middle childhood: A short-term longitudinal study. *Journal of Experimental Child Psychology, 163*, 69-86. doi:10.1016/j.jecp.2017.06.011

Lecce, S., Bianco, F., Devine, R. T., Hughes, C., & Banerjee, R. (2014). Promoting theory of mind during middle childhood: A training program. *Journal of Experimental Child Psychology, 126*, 52-67. doi:10.1016/j.jecp.2014.03.002

Lecce, S., Bianco, F., & Ronchi, L. (2020). Executive function in the school context: The role of peer relationships. *Infant and Child Development, 29*(1), e2151. doi:10.1002/icd.2151

Lecce, S., Caputi, M., & Hughes, C. (2011). Does sensitivity to criticism mediate the relationship between theory of mind and academic achievement? *Journal of Experimental Child Psychology, 110*(3), 313-331. doi:10.1016/j.jecp.2011.04.011

Lecce, S., Zocchi, S., Pagnin, A., Palladino, P., & Taumoepeau, M. (2010). Reading minds: the relation between children's mental state knowledge and their metaknowledge about reading. *Child Development, 81*(6), 1876-1893. doi:10.1111/j.1467-8624.2010.01516.x

Leonard, M. A., Milich, R., & Lorch, E. P. (2011). The Role of Pragmatic Language Use in Mediating the Relation Between Hyperactivity and Inattention and Social Skills Problems. *Journal of Speech Language and Hearing Research, 54*(2), 567. doi:10.1044/1092-4388(2010/10-0058)

Loukusa, S., & Leinonen, E. (2008). Development of comprehension of ironic utterances in 3- to 9-year-old Finnish-speaking children. *Psychology of Language and Communication, 12*(1), 55-69. doi:10.2478/v10057-008-0003-0

Marconi, L., Ott, M., Pesenti, E., Ratti, D., & Tavella, M. (1994). *Lessico Elementare. Dati statistici sull'italiano scritto e letto dai bambini delle elementari.* Bologna, Italy: Zanichelli.

Markman, E. M., & Wachtel, G. F. (1988). Children's use of mutual exclusivity to constrain the meanings of words. *Cognitive Psychology, 20*(2), 121-157. doi:10.1016/0010-0285(88)90017-5

Markman, E. M., Wasow, J. L., & Hansen, M. B. (2003). Use of the mutual exclusivity assumption by young word learners. *Cognitive Psychology, 47*(3), 241-275. doi:10.1016/S0010-0285(03)00034-3

Marks, P. E. L., Babcock, B., Cillessen, A. H. N., & Crick, N. R. (2013). The Effects of Participation Rate on the Internal Reliability of Peer Nomination Measures. *Social Development, 22*(3), 609-622. doi:10.1111/j.1467-9507.2012.00661.x

Martin, I., & McDonald, S. (2005). Evaluating the causes of impaired irony comprehension following traumatic brain injury. *Aphasiology, 19*(8), 712-730. doi:10.1080/02687030500172203

Mashal, N., & Kasirer, A. (2011). Thinking maps enhance metaphoric competence in children with autism and learning disabilities. *Research in Developmental Disabilities, 32*(6), 2045-2054. doi:10.1016/j.ridd.2011.08.012

Massaro, D., Valle, A., & Marchetti, A. (2014). Do social norms, false belief understanding, and metacognitive vocabulary influence irony comprehension? A study of five- and seven-year-old children. *European Journal of Developmental Psychology, 11*(3), 292-304. doi:10.1080/17405629.2013.821407

Masten, A. S., Roisman, G. I., Long, J. D., Burt, K. B., Obradović, J., Riley, J. R., . . . Tellegen, A. (2005). Developmental cascades: linking academic achievement and externalizing and internalizing symptoms over 20 years. *Developmental Psychology, 41*(5), 733-746. doi:10.1037/0012-1649.41.5.733

Matthews, D., Biney, H., & Abbot-Smith, K. (2018). Individual Differences in Children's Pragmatic Ability: A Review of Associations with Formal Language, Social Cognition, and Executive Functions. *Language Learning and Development, 14*(3), 186-223. doi:10.1080/15475441.2018.1455584

Matthews, D., Lieven, E., & Tomasello, M. (2010). What's in a manner of speaking? Children's sensitivity to partner-specific referential precedents. *Developmental Psychology, 46*(4), 749-760. doi:10.1037/a0019657

Melogno, S., Pinto, M. A., & Di Filippo, G. (2017). Sensory and Physico-Psychological Metaphor Comprehension in Children with ASD: A Preliminary Study on the Outcomes of a Treatment. *Brain Sci, 7*(7). doi:10.3390/brainsci7070085

Melot, A. M., & Angeard, N. (2003). Theory of mind: is training contagious? *Developmental Science, 6*(2), 178-184. doi:10.1111/1467-7687.00269

Milligan, K., Astington, J. W., & Dack, L. A. (2007). Language and theory of mind: meta-analysis of the relation between language ability and false-belief understanding. *Child Development, 78*(2), 622-646. doi:10.1111/j.1467-8624.2007.01018.x

Miranda, A., Berenguer, C., Roselló, B., & Baixauli, I. (2020). Relationships between the social communication questionnaire and pragmatic language, socialization skills, and behavioral problems in children with autism spectrum disorders. *Applied Neuropsychology: Child,* 1–12. doi:10.1080/21622965.2018.1550403

Mo, S., Su, Y., Chan, R. C., & Liu, J. (2008). Comprehension of metaphor and irony in schizophrenia during remission: the role of theory of mind and IQ. *Psychiatry Res, 157*(1-3), 21-29. doi:10.1016/j.psychres.2006.04.002

Montemurro, S., Mondini, S., Signorini, M., Marchetto, A., Bambini, V., & Arcara, G. (2019). Pragmatic Language Disorder in Parkinson's Disease

and the Potential Effect of Cognitive Reserve. *Frontiers in Psychology, 10*(1220). doi:10.3389/fpsyg.2019.01220

Muraven, M., & Baumeister, R. F. (2000). Self-regulation and depletion of limited resources: Does self-control resemble a muscle? *Psychological Bulletin, 126*(2), 247-259. doi:10.1037/0033-2909.126.2.247

Muraven, M., Tice, D. M., & Baumeister, R. F. (1998). Self-control as a limited resource: Regulatory depletion patterns. *Journal of Personality and Social Psychology, 74*(3), 774-789. doi:10.1037/0022-3514.74.3.774

Nelson, K. (2007). *Young minds in social worlds: Experience, meaning, and memory*. Cambridge, MA, US: Harvard University Press.

Newcomb, A. F., Bukowski, W. M., & Pattee, L. (1993). Children's peer relations: a meta-analytic review of popular, rejected, neglected, controversial, and average sociometric status. *Psychological Bulletin, 113*(1), 99-128. doi:10.1037/0033-2909.113.1.99

Nilsen, E.S., Glenwright, M., & Huyder, V. (2011) Children and Adults Understand That Verbal Irony Interpretation Depends on Listener Knowledge, *Journal of Cognition and Development, 12*(3), 374-409. doi:10.1080/15248372.2010.544693

Nippold, M. A., Leonard, L. B., & Kai, l. R. (1984). Syntactic and Conceptual Factors in Children's Understanding of Metaphors. *Journal of Speech, Language, and Hearing Research, 27*(2), 197-205. doi:doi:10.1044/jshr.2702.197

Norbury, C. F. (2005). The relationship between theory of mind and metaphor: Evidence from children with language impairment and autistic spectrum disorder. *British Journal of Developmental Psychology, 23*(3), 383-399. doi:10.1348/026151005x26732

Noveck, I. A., Bianco, M., & Castry, A. (2001). The Costs and Benefits of Metaphor. *Metaphor and Symbol, 16*(1-2), 109-121. doi:10.1080/10926488.2001.9678889

Nowicki, S., & Oxenford, C. (1989). The Relation of Hostile Nonverbal Communication Styles to Popularity in Preadolescent Children. *Journal of Genetic Psychology, 150*(1), 39–43. doi:10.1080/00221325.1989.9914572

O'Neill, D. K. (1996). Two-Year-Old Children's Sensitivity to a Parent's Knowledge State When Making Requests. *Child Development, 67*(2), 659-677. doi:10.2307/1131839

O'Neill, D. K., & Topolovec, J. C. (2001). Two-year-old children's sensitivity to the referential (in)efficacy of their own pointing gestures. *Journal of Child Language, 28*(1), 1-28. doi:10.1017/S0305000900004566

Oberle, E., & Schonert-Reichl, K. A. (2013). Relations among peer acceptance, inhibitory control, and math achievement in early adolescence. Journal of *Applied Developmental Psychology, 34*(1), 45–51. doi:10.1016/j.appdev.2012.09.003

Olson, D. R., Antonietti, A., Liverta-Sempio, O., & Marchetti, A. (2006). The Mental Verbs in Different Conceptual Domains and in Different Cultures. In A. Antonietti, O. Liverta-Sempio, & A. Marchetti (Eds.), *Theory of*

Mind and Language in Developmental Contexts (pp. 31-64). Boston, MA: Springer US.

Ornaghi, V., Brockmeier, J., & Gavazzi, I. G. (2011). The Role of Language Games in Children's Understanding of Mental States: A Training Study. *Journal of Cognition and Development, 12*(2), 239-259. doi:10.1080/15248372.2011.563487

Ornaghi, V., & Grazzani, I. (2013). The relationship between emotional-state language and emotion understanding: a study with school-age children. *Cognition and Emotion, 27*(2), 356-366. doi:10.1080/02699931.2012.711745

Orsini, A. (1993). *WISC-R. Contributo alla taratura italiana*. Florence, Italy: Organizzazioni Speciali.

Ortony, A. (1975). Why Metaphors Are Necessary and Not Just Nice1. *Educational Theory, 25*(1), 45-53. doi:10.1111/j.1741-5446.1975.tb00666.x

Perner, J., & Wimmer, H. (1985). "John thinks that Mary thinks that…": Attribution of second-order beliefs by 5- to 10-year-old children. *Journal of Experimental Child Psychology, 39*(3), 437-471. doi:10.1016/0022-0965(85)90051-7

Persicke, A., Tarbox, J., Ranick, J., & St. Clair, M. (2012). Establishing metaphorical reasoning in children with autism. *Research in Autism Spectrum Disorders, 6*(2), 913-920. doi:10.1016/j.rasd.2011.12.007

Peterson, C. C., & Slaughter, V. (2003). Opening windows into the mind: Mothers' preferences for mental state explanations and children's theory of mind. *Cognitive Development, 18*(3), 399-429. doi:10.1016/S0885-2014(03)00041-8

Petranovich, C. L., Walz, N. C., Staat, M. A., Chiu, C.-Y. P., & Wade, S. L. (2017). Structural language, pragmatic communication, behavior, and social competence in children adopted internationally: A pilot study. *Applied Neuropsychology: Child, 6*(4), 315–326. doi:10.1080/21622965.2016.1182433

Philofsky, A., Fidler, D. J., & Hepburn, S. (2007). Pragmatic language profiles of school-age children with autism spectrum disorders and Williams syndrome. *American Journal of Speech Language Pathology, 16*(4), 368-380. doi:10.1044/1058-0360(2007/040)

Piaget, J. (1932). *The moral judgment of the child.* Oxford, England: Harcourt, Brace.

Pinto, M. A., Melogno, S., & Iliceto, P. (2006). *TCM. Test di comprensione delle metafore. Scuola elementare e scuola media*: Carocci Faber.

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

Pinto, M. A., Melogno, S., & Iliceto, P. (2011). Assessing metaphor comprehension as a metasemantic ability in students from 9 to 14 years-old. *Linguarum Arena*, 57-77.

Place, K. S., & Becker, J. A. (1991). The influence of pragmatic competence on the likeability of grade-school children. *Discourse Processes, 14*(2), 227-241. doi:10.1080/01638539109544783

Pouscoulous, N. (2014). "The elevator's buttocks": Metaphorical abilities in children. In D. Matthews (Ed.), *Pragmatic development in first language acquisition* (Vol. Trends in language acquisition research). Amsterdam: John Benjamins Publishing Company.

Putallaz, M., & Gottman, J. M. (1981). An Interactional Model of Children's Entry into Peer Groups. *Child Development, 52*(3), 986-994. doi:10.2307/1129103

Putallaz, M., & Wasserman, A. (1989). Children's naturalistic entry behavior and sociometric status: A developmental perspective. *Developmental Psychology, 25*(2), 297-305. doi:10.1037/0012-1649.25.2.297

Rabiner, D. L., & Gordon, L. V. (1992). The Coordination of Conflicting Social Goals: Differences between Rejected and Nonrejected Boys. *Child Development, 63*(6), 1344-1350. doi:10.2307/1131560

Ronchi, L., Banerjee, R., & Lecce, S. (2020). Theory of mind and peer relationships: The role of social anxiety. *Social Development, 29*(2), 478-493. doi:10.1111/sode.12417

Rubin, K. H., Bukowski, W. M., & Parker, J. G. (2007). Peer Interactions, Relationships, and Groups *Handbook of child psychology: Social, emotional, and personality development, Vol. 3, 6th ed.* (pp. 571-645). Hoboken, NJ, US: John Wiley & Sons Inc.

Rubini, V., & Rossi, M. A. (1982). Analysis of the psychometric characteristics of the Primary Mental Abilities Tests (Level 4-sup-6, 1962 revision) applied to a group of Italian subjects. *Bollettino di Psicologia Applicata, 161*, 87-105. doi:10.5840/newvico1984241

Ruffman, T., Slade, L., & Crowe, E. (2002). The Relation between Children's and Mothers' Mental State Language and Theory-of-Mind Understanding. *Child Development, 73*(3), 734-751. doi:10.1111/1467-8624.00435

Rundblad, G., & Annaz, D. (2010). Development of metaphor and metonymy comprehension: receptive vocabulary and conceptual knowledge. *British Journal of Developmental Psychology, 28*(Pt 3), 547-563. doi:10.1348/026151009X454373

Satorra, A. (2000). Scaled and Adjusted Restricted Tests in Multi-Sample Analysis of Moment Structures. In R. D. H. Heijmans, D. S. G. Pollock, & A. Satorra (Eds.), *Innovations in Multivariate Statistical Analysis: A Festschrift for Heinz Neudecker* (pp. 233-247). Boston, MA: Springer US.

Satorra, A., & Bentler, P. M. (2001). A scaled difference chi-square test statistic for moment structure analysis. *Psychometrika, 66*(4), 507-514. doi:10.1007/BF02296192

Sawyer, R. K. (1997). *Pretend play as improvisation: Conversation in the preschool classroom*. Hillsdale, NJ, US: Lawrence Erlbaum Associates, Inc.

Searle, J. R. (1979). What is an intentional state? *Mind, 88*(January), 74-92. doi:10.1093/mind/LXXXVIII.1.74

Seigneuric, A., Megherbi, H., Bueno, S., Lebahar, J., & Bianco, M. (2016). Children's comprehension skill and the understanding of nominal metaphors. *Journal of Experimental Child Psychology, 150*, 346-363. doi:10.1016/j.jecp.2016.06.008

Selig, J. P., & Little, T. D. (2012). Autoregressive and cross-lagged panel analysis for longitudinal data. In B. Laursen, T. D. Little, & N. A. Card (Eds.), *Handbook of developmental research methods* (pp. 265-278). New York, NY, US: The Guilford Press

Senju, A., & Csibra, G. (2008). Gaze Following in Human Infants Depends on Communicative Signals. *Current Biology, 18*(9), 668-671. doi:10.1016/j.cub.2008.03.059

Serratrice, L. (2005). The role of discourse pragmatics in the acquisition of subjects in Italian. *Applied Psycholinguistics*, 26(03), 437 - 462. doi: 10.1017/S0142716405050241

Shatz, M., & Gelman, R. (1973). The development of communication skills: modifications in the speech of young children as a function of listener. *Monographs of the Society for Research in Child Development, 38*(5), 1-38.

Slaughter, V., & Gopnik, A. (1996). Conceptual Coherence in the Child's Theory of Mind: Training Children to Understand Belief. *Child Development, 67*(6), 2967-2988. doi:10.2307/1131762

Slaughter, V., Imuta, K., Peterson, C. C., & Henry, J. D. (2015). Meta-Analysis of Theory of Mind and Peer Popularity in the Preschool and Early School Years. *Child Development, 86*(4), 1159–1174. doi:10.1111/cdev.12372

Sopory, P., & Dillard, J. P. (2002). The persuasive effects of metaphor: A meta-analysis. *Human Communication Research, 28*(3), 382-419. doi:10.1111/j.1468-2958.2002.tb00813.x

Sperber, D., & Wilson, D. (1995). *Relevance: Communication and cognition*. Blackwell, Oxford, United Kingdom.

Sperber, D., & Wilson, D. (2002). Pragmatics, Modularity and Mind-reading. *Mind & Language, 17(1-2)*, 3-23. doi:10.1111/1468-0017.00186

Sperber, D., Wilson, D., 2005. *Pragmatics.* In: F. Jackson & M. Smith (Eds.), Oxford handbook of contemporary philosophy. Oxford University Press, Oxford, United Kingdom.

St Clair, M. C., Pickles, A., Durkin, K., & Conti-Ramsden, G. (2011). A longitudinal study of behavioral, emotional and social difficulties in individuals with a history of specific language impairment (SLI). *J Commun Disord, 44*(2), 186-199. doi:10.1016/j.jcomdis.2010.09.004

Stella, G., Pizzioli, C., & Tressoldi, P. (2000). *Peabody Picture Vocabulary Test-revised (PPVT-R). Adattamento italiano e standardizzazione.* Turin, Italy: Omega Edizioni.

Suraniti, S., Ferri, R., & Neri, V. (2009). *Test for Reception of Grammar version 2 (TROG-2)*. Florence, Italy: OS.

Szufnarowska, J., Rohlfing, K. J., Fawcett, C., & Gredebäck, G. (2014). Is ostension any more than attention? *Scientific Report, 4*(1), 5304. doi:10.1038/srep05304

Thurstone, L. L., & Thurstone, T. G. (1962). *Primary mental abilities.* Chicago, Illinois: Science Research.

Tonini, E., Lecce, S., Del Sette, P., Bianco, F., Canal, P., Bambini, V. (2020). *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. doi:10.1159/000277297

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. doi:10.1177/0165025412463508

van der 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. doi:10.1016/j.cedpsych.2018.06.007

van der Wilt, F., van der Veen, C., van Kruistum, C., & van Oers, B. (2019). Why Do Children Become Rejected by Their Peers? A Review of Studies into the Relationship Between Oral Communicative Competence and

Sociometric Status in Childhood. *Educational Psychology Review, 31*(3), 699-724. doi:10.1007/s10648-019-09479-z

van der Wilt, F., van der Veen, C., van Kruistum, C., & van Oers, B. (2020). Language Abilities and Peer Rejection in Kindergarten: A Mediation Analysis. *Early Education and Development, 31*(2), 308-322. doi:10.1080/10409289.2019.1624145

van der Wilt, F., van Kruistum, C., van der Veen, C., & van Oers, B. (2016). Gender differences in the relationship between oral communicative competence and peer rejection: an explorative study in early childhood education. *European Early Childhood Education Research Journal, 24*(6), 807-817. doi:10.1080/1350293X.2015.1073507

Volden, J., Coolican, J., Garon, N., White, J., & Bryson, S. (2008). Brief Report: Pragmatic Language in Autism Spectrum Disorder: Relationships to Measures of Ability and Disability. *Journal of Autism and Developmental Disord*ers, 39(2), 388-393. doi:10.1007/s10803-008-0618-y

Vosniadou, S., Ortony, A., Reynolds, R. E., & Wilson, P. T. (1984). Sources of difficulty in the young child's understanding of metaphorical language. *Child Development*, 1588-1606. doi:10.2307/1130028

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

Vygotsky, L. S., Cole, M., John-Steiner, V., Scribner, S., & Souberman, E. (1978). *Mind in Society: Development of Higher Psychological Processes*: Harvard University Press.

Wang, C., & Dowker, A. (2010). A cross-cultural study of metaphoric understanding. In A. Deignan & L. Cameron (Eds.), *Researching and applying metaphor in the real world* (pp. 105–122). Amsterdam: John Benjamins.

Watson, A. C., Nixon, C. L., Wilson, A., & Capage, L. (1999). Social interaction skills and theory of mind in young children. *Developmental Psychology, 35*(2), 386–391. doi:10.1037/0012-1649.35.2.386

Wellman, H. M., Cross, D., & Watson, J. (2001). Meta-Analysis of Theory-of-Mind Development: The Truth about False Belief. *Child Development, 72*(3), 655-684. doi:10.1111/1467-8624.00304

Wellman, H. M., & Liu, D. (2004). Scaling of Theory-of-Mind Tasks. *Child Development, 75*(2), 523-541. doi:10.1111/j.1467-8624.2004.00691.x

Wentzel, K. R., & Caldwell, K. (1997). Friendships, Peer Acceptance, and Group Membership: Realtions to Academic Achievement in Middle School. *Child Development, 68*(6), 1198–1209. doi:10.1111/j.1467-8624.1997.tb01994.x

White, S., Hill, E., Happe, F., & Frith, U. (2009). Revisiting the strange stories: revealing mentalizing impairments in autism. *Child Development, 80*(4), 1097-1117. doi:10.1111/j.1467-8624.2009.01319.x

Whyte, E. M., & Nelson, K. E. (2015). Trajectories of pragmatic and nonliteral language development in children with autism spectrum disorders. *J Commun Disord, 54*, 2-14. doi:10.1016/j.jcomdis.2015.01.001

Wilson, D., & Carston, R. (2007). *A unitary approach to lexical pragmatics: Relevance, inference and ad hoc concepts*: na.

Wimmer, H., & Perner, J. (1983). Beliefs about beliefs: representation and constraining function of wrong beliefs in young children's understanding of deception. *Cognition, 13*(1), 103-128. doi:10.1016/0010-0277(83)90004-5

Winner, E., Engel, M., & Gardner, H. (1980). Misunderstanding metaphor: what's the problem? *Journal of Experimental Child Psychology, 30*(1), 22-32. doi:10.1016/0022-0965(80)90072-7

Winner, E., Levy, J., Kaplan, J., & Rosenblatt, E. (1988). Children's Understanding of Nonliteral Language. *Journal of Aesthetic Education, 22*(1), 51-63. doi:10.2307/3332964

Winner, E., Rosenstiel, A. K., & Gardner, H. (1976). The development of metaphoric understanding. *Developmental Psychology, 12*(4), 289-297. doi:10.1037/0012-1649.12.4.289

Zarbatany, L., Hartmann, D. P., & Rankin, D. B. (1990). The psychological functions of preadolescent peer activities. *Child Development, 61*(4), 1067-1080.

Zufferey, S. (2010). *Lexical Pragmatics and Theory of Mind: The Acquisition of Connectives*. Amsterdam, The Netherlands: John Benjamins.

Appendices

Appendix A. Tasks adopted in the present dissertation. For each task is reported studies in which were used and chapters in which were described.

Construct	Tasks	Study	Chapter
Socio-economic status	Family Affluence Scale (FAS; Currie et al., 2008)	Study 1 Study 2	Chapter 1 Chapter 2
Verbal abilities	Primary Mental Abilities (PMA; Rubini & Rossi, 1982; Thurstone & Thurstone, 1962); vocabulary subtest, intermediate form	Study 1	Chapter 1
	Primary Mental Abilities (PMA; Rubini & Rossi, 1982; Thurstone & Thurstone, 1962); vocabulary subtest	Study 2 Study 3	Chapter 2 Chapter 3
	Peabody Picture Vocabulary Test–Revised (Dunn & Dunn, 1981)	Study 4	Chapter 4
	Test for Reception of Grammar (Bishop, 2003)	Study 4	Chapter 4
Reading comprehension	MT task (Cornoldi & Colpo, 1998)	Study 4	Chapter 4
Working memory	Backward Digit Span task of the WISC-R Italian version (Orsini, 1993)	Study 2 Study 4	Chapter 2 Chapter 4
Inferential skills	Strange Stories task (Happé, 1994); physical-control stories	Study 2 Study 3 Study 4	Chapter 2 Chapter 3 Chapter 4
Theory of Mind	Strange Stories task (Happé, 1994); mental stories	Study 2 Study 3 Study 4	Chapter 2 Chapter 3 Chapter 4
	Animated triangles tasks (Castelli, Happé, Frith, & Frith, 2000)	Study 4	Chapter 4
Peer relationships	Sociometric peer nominations (Coie, Dodge, & Coppotelli; 1982); most-like and least-like	Study 1	Chapter 1
Metaphor Understanding	Physical and Mental Metaphors task – extended version	Study 1 Study 3	Chapter 1 Chapter 3
	Physical and Mental Metaphors task – short version	Study 2	Chapter 2
	Physical and Mental Metaphors task – hard version	Study 4	Chapter 4
	Referential task (Noveck, Bianco, & Castry, 2001)	Study 4	Chapter 4

Appendix B. Example of a story and exercises of the ToM training (Chapter 4).

Maria really likes chocolate snacks. One of her classmates, Alice, wants to play a joke on her. So, while Maria is in the garden during the break, Alice hides Maria's snack in her case. When Maria goes back to the class, she doesn't find her snack, but she sees Alice laughing. Even though Maria believes that Alice will lie to her, she asks her: "Where did you hide my snack? You must have put it in your pencil case or in your backpack, because I don't find it in other places. So where is my snack? In the case or in the backpack?". Alice answers: "I put it in my backpack!".

Questions and exercises:

1. Where is the snack?
2. Why Alice answers that the snack is in her backpack?
3. Do you think that Alice will be able to deceive Maria by saying that the snack is in the backpack?
4. Can you imagine a better way to fool Maria? What should Alice say? Why? (Remember that Maria can't look in both places, and that she believes Alice will probably lie to her).
5. What is, in your opinion, the meaning of this sentence "Maria *believes* that Alice will lie to her"
    - Maria *knows* that Alice will lie to her
    - Maria *remembers* that Alice will lie to her
    - Maria *decides* that Alice will lie to her

- Maria *thinks* that Alice will lie to her

Appendix C. Example of a story and exercises of the MetaCoM training (Chapter 4).

During the process, the thief explains that he attempted to rob the shop because he had no money to feed his children after he was fired. Despite this, Judge Santino condemns him to the full sentence: ten years in prison. The poor man's lawyer says: "Judge Santino is really a computer, he could have saved him at least a few years".

Questions and exercises:

1. What Judge Santino and a computer have in common?
2. In the story, the lawyer says that Judge Santino is a computer. What does it mean?
    - That Judge Santino is very good at pronouncing sentences.
    - That Judge Santino has the keyboard and the mouse.
    - That Judge Santino is insensitive to the thief's situation.
3. Why does the lawyer say that?
4. Try to create some metaphors. For example you can choose the words cheeks and apples because cheeks are red and round like apples. Now try with the following words:

    computer – enemy – keyboard – water
5. Have you ever used the computer metaphor to talk about someone who is severe, cold and does not take pity? If yes, tell us. Otherwise try to create a story where the characters use this metaphor.

Acknowledgements

Firstly, I would like to acknowledge that the present dissertation and the studies inside it were 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.

I would like to start the acknowledgments by thanking my supervisor Prof. Serena Lecce for inspiring and encouraging me during my travel in the field of developmental scientific research. I am also grateful to Prof. Valentina Bambini who follows this project and shares with me her knowledge and experience in the field of pragmatics. Finally, I would like to express my gratitude to my colleagues, my friends, my family, and my husband for walking with me throughout my PhD experience.