Cognitive Development 57 (2021) 101003

Contents lists available at ScienceDirect

Cognitive Development

journal homepage: www.elsevier.com/locate/cogdev

Interactional training interventions boost children's expressive pragmatic abilities: evidence from a novel multidimensional testing approach

Mariia Pronina [a,*], Iris Hübscher [b], Judith Holler [c,d], Pilar Prieto [e,a]

[a] Department of Translation and Language Sciences, Universitat Pompeu Fabra, Barcelona, Catalonia, Spain
[b] URPP Language and Space, University of Zurich, Switzerland
[c] Donders Institute for Brain, Cognition and Behaviour, Radboud University, Nijmegen, the Netherlands
[d] Max Planck Institute for Psycholinguistic, Nijmegen, the Netherlands
[e] Institució Catalana de Recerca i Estudis Avançats (ICREA), Barcelona, Catalonia, Spain

ARTICLE INFO

Keywords:
Pragmatics
Theory of mind
Mental states
Conversational intervention
Embodiment
Multimodal language

ABSTRACT

This study investigates the effectiveness of training preschoolers in order to enhance their social cognition and pragmatic skills. Eighty-three 3–4-year-olds were divided into three groups and listened to stories enriched with mental state terms. Then, whereas the control group engaged in non-reflective activities, the two experimental groups were guided by a trainer to reflect on mental states depicted in the stories. In one of these groups, the children were prompted to not only talk about these states but also "embody" them through prosodic and gestural cues. Results showed that while there were no significant effects on Theory of Mind, emotion understanding, and mental state verb comprehension, the experimental groups significantly improved their pragmatic skill scores pretest-to-posttest. These results suggest that interactional interventions can contribute to preschoolers' pragmatic development, demonstrate the value of the new embodied training, and highlight the importance of multidimensional testing for the evaluation of intervention effects.

1. Introduction

In the preschool years, children gradually develop the ability to interact with the social world around them. This ability is highly relevant for children's everyday life because it impacts their social relationships and the quality of their interactions (Matthews, Biney, & Abbot-Smith, 2018). Much attention has been devoted to studying the ability to interact socially and its underlying skills, particularly in the domain of social cognition and pragmatics. The present article builds on this previous research and examines how children's developing social cognition and pragmatic abilities can be boosted in the preschool years by focused training.

Social cognition is an umbrella term that refers to a set of cognitive and emotional abilities that are applied in social situations (Harvey & Penn, 2010). One ability generally included in this set is the ability to attribute mental states to other human beings and to make predictions about their behavior, known as Theory of Mind (ToM), a term coined by Premack and Woodruff (1978); see also Perner, 1991, and Wellman, 2014, for a review). In the developmental field, ToM skills have received considerable attention,

* Corresponding author at: Department of Translation and Language Sciences - Universitat Pompeu Fabra, Roc Boronat, 138, 08018, Barcelona, Catalonia, Spain.
E-mail address: mariia.pronina@upf.edu (M. Pronina).

https://doi.org/10.1016/j.cogdev.2020.101003
Received 31 July 2020; Received in revised form 13 November 2020; Accepted 21 December 2020
Available online 16 January 2021
0885-2014/© 2021 Elsevier Inc. All rights reserved.

particularly in the last forty years (see Henry, Phillips, Ruffman, & Bailey, 2013; Wellman, 2018, for a review). A large number of studies have demonstrated that this developmental milestone is reached in the preschool years, as children gradually develop the ability to detect and interpret the mental states of themselves and others (see Perner & Roessler, 2012; Wellman, Cross, & Watson, 2001; Wellman & Liu, 2004).

At present, ToM in children is traditionally measured by means of what are known as false belief tests (Wimmer & Perner, 1983). In order to succeed in these tests, the child must be able to reconstruct the false belief of others or his/her own previous false belief. For example, in the so-called "Sally and Anne" change-of-location task (Baron-Cohen, Leslie, & Frith, 1985), the child is required to predict that a puppet called Sally will mistakenly look for an object, which has been moved to a different place during her absence by Anne, another puppet, in the place where she last left it rather than in the place to which Anne has moved it. Another widely used standard false belief test is the "Smarties" unexpected content task (Astington & Gopnik, 1988; Perner, Leekam, & Wimmer, 1987) in which children are asked to recall their initial false belief about the contents of a Smarties candy tube. However, just using false belief tests may not be sufficient to assess a child's developing social cognition, as it provides us only with a small glimpse of what social cognition is about (see Harris, 2006, for an overview of children's developing understanding of belief in the context of social cognition; Wellman, 2018, for a criticism of the exclusive use of the false belief paradigm). Another important aspect of social cognition is emotion understanding, which refers to the ability to infer others' emotions and feelings. It has been shown that even at 3.5 months of age, infants can recognize certain emotions in others through facial and prosodic cues (Kahana-Kalman & Walker-Andrews, 2001). However, it is only when they approach school age—as with their ability to detect false beliefs—that they become able to name emotions and understand the external causes of emotions in others (Pons, Harris, & de Rosnay, 2004).

As stated, the general focus of the literature on social cognition development has been on ToM skills, and, particularly, on the ability to understand false belief (see Harris, 2006 for an overview of children's developing understanding of belief in the context of social cognition; Wellman, 2018 for a criticism of the exclusive use of the false belief paradigm), with other components such as pragmatic skills receiving less attention. Pragmatic skills usually include the ability to recognize and express communicative intentions, infer other people's expectations and motivations, initiate conversations, respond with relevant information, and understand utterances in context. The development of these pragmatic skills, which are crucial in a child's daily social life (Matthews, 2014; Matthews et al., 2018; Norbury, 2014), is characterized by its own trajectory (Domaneschi & Bambini, 2020). However, despite its overall importance and relation to social cognition, the ability to socially interact with others has not been subjected to much research in the domain of social cognition (see Kunda, 1999, for the range of phenomena attributed to social cognition).

In order to address a child's developing social ability more comprehensively, it might be beneficial to focus on several skills simultaneously. In the present study we will undertake to do that by looking at the impact on preschool children of a class-room based intervention in terms of not only false belief understanding but also other aspects of social cognition (cognitive and affective aspects), as well as pragmatic skills. Such skills help children understand social behavior and social interactions and are closely linked to language development (Conte, Ornaghi, Grazzani, Pepe, & Cavioni, 2019; Ornaghi, Pepe, & Grazzani, 2016). Both are also related to social competence (see Hughes & Leekam, 2004) and impact children's developing social relationships, influencing, for example, their popularity with peers, their capacity for leadership (Peterson, O'Reilly, & Wellman, 2016), and even their academic achievement (Lecce, Caputi, & Hughes, 2011; Lecce, Caputi, & Pagnin, 2014).

We hypothesize that training children to talk about mental states could potentially enhance not only their ToM understanding but their pragmatic competence behavior overall. Boosting the whole set of socio-communicative abilities early on in a child's development may also be important from an educational point of view.

1.1. ToM training

In recent decades, training studies have focused on children's ToM development and have shown the beneficial effects of focused language-based interventions on ToM. The first ToM training studies focused on the use of false belief tasks and demonstrated that a child's performance on these tasks can be improved if the child is exposed to mental state language (Appleton & Reddy, 1996; Melot & Angeard, 2003; Slaughter, 1998; Slaughter & Gopnik, 1996). These studies used a task specific approach, namely tasks which were specifically developed to help preschool children pass standard false belief tests, which generally consisted of providing them with feedback right after their performance on specific false belief sequences. For example, in Appleton and Reddy (1996), after two weeks of training in groups, 3-year-old children showed improvement in their ability to pass false belief tasks about the location of an object. Positive results were also reported by Slaughter and Gopnik (1996), who individually trained 3- to 4-year-olds on appearance vs. reality distinctions for two weeks, and by Slaughter (1998), who carried out individual training sessions on false belief unexpected content tasks with 3- to 4-year-olds over the course of two or three weeks. In Clements, Rustin, and McCallum (2000), 3- to 5-year-old children were individually trained to pass change-of-location false belief tasks by being provided with detailed explanations of their judgements, which was also found to be effective. In another successful two-week intervention study by Melot and Angeard (2003), 3- to 4-year-old children were provided with explicit feedback on their performance in change-of-location false belief tasks and appearance vs. reality distinction tasks. A more recent study by Lecce, Bianco, Demicheli, and Cavallini (2014) adopted this same approach and similarly showed that training 4- to 5-year olds individually in three 20-minute sessions improved their performance on false belief unexpected content and unexpected location tasks.

While these first studies showed that specific training providing feedback and explanations about the correctness of children's answers in false belief tasks can lead to improvements in ToM, other studies have tested whether training linguistic skills (specifically syntactic skills) can facilitate the development of ToM (Hale & Tager-Flusberg, 2005; Lohmann & Tomasello, 2003). A close relationship between ToM development and language has been shown in research (Astington & Jenkins, 1999; de Villiers, 1995, 2000;

Happé, 1995; more recently Milligan, Astington, & Dack, 2007; Tager-Flusberg & Joseph, 2012), particularly the link between ToM and the acquisition of sentential complement syntax (de Villiers & Pyers, 2002; Slade & Ruffman, 2005). According to de Villier's account (de Villiers, 1995, 2000; de Villiers & De Villiers, 2000), the acquisition of sentential complements promotes false belief understanding because it provides children with the representational format needed for reasoning about false beliefs. More precisely, since in sentential complements the clause that may be false can be embedded in a main clause that is true, the very structure of complement clauses makes it possible to handle misrepresentation and serves as a bootstrap for false belief understanding. Based on this, Lohmann and Tomasello (2003) conducted a two-week training intervention with 3-year-old children that focused on discourse-shifting (shifting between different perspectives while talking about deceptive objects) and sentential complement syntax. The results showed that language-based training played a key role in the development of the children's false belief understanding. In a similar study, Hale and Tager-Flusberg (2003) trained children either in sentential complements or, as a control condition, in relative clauses. While the control group showed no improvement in false belief understanding at posttest, the group trained in sentential complements improved their performance on a false belief task. Moreover, training children in sentential constructions had the same effect on false belief understanding as false belief training per se.

Another set of training studies have used conversation-based approaches to boost ToM performance (Astington & Peskin, 2004; Esteban, Sidera, Serrano, Amadó, & Rostan, 2010; Guajardo & Watson, 2002; Ornaghi, Brockmeier, & Grazzani, 2011). All of these studies have used storybooks whose texts included a variety of terms describing mental states[1] such as *know*, *think*, *get angry*, *get scared* to train ToM. However, contradictory results have been obtained. For example, Guajardo and Watson (2002) examined in two studies whether exposing children to storybooks which focused on mental states would facilitate ToM understanding. In Study 1, a group of 37 3- to 4-year-old children participated in 12–15 training sessions lasting 15–25 min for five weeks which involved an experimenter reading storybooks in class and subsequently discussing mental state concepts with the child participants. The training did not lead to significant improvement in any of the false belief (unexpected location and unexpected content) posttest tasks. In Study 2, by contrast, which involved 54 slightly younger children, training was given on an individual basis rather than in a group and sessions lasted 10–15 min. In this case, training was found to be effective. A separate training study by Astington and Peskin (2004) yielded negative findings. In this four-week study, storybooks were read to a total of 48 4-year-old children at home, individually, and at school, as a group activity. However, while one group (the experimental group) was exposed to texts with a large number of mental state terms, the other (the control group) were told the same stories but without any explicit mental state terms. Though both groups showed improvement, surprisingly, children in the control group improved significantly more in a subsequent false belief explanation task than children in the experimental group. Another study by Esteban et al. (2010) with 96 3- to 4-year-old children also looked at the impact of storybook reading and subsequent discussion on children's performance on several individual ToM tasks. In this case, the experimental group of child participants was first told a story by their teacher and then participated in conversations and reflective interactions about the story. This training session lasted 15 min and took place three times over two weeks. By contrast, the control group was exposed to no special intervention. The results showed that the experimental group improved only in an unexpected location task, while no effect was found for an unexpected content task. In a similar study involving two age groups (34 3-year-olds and 36 4-year--olds), Ornaghi et al. (2011) explored the role of actively using mental state lexical items in conversational activities based on a storybook in promoting performance on ToM tasks. The time frame for training was much longer in this study, since the 20-minute sessions took place twice a week over two months. In the experimental condition, children in groups of six were exposed to one of the stories from the storybook and afterwards engaged in language games and conversations which encouraged them to use mental state terms. In the control condition, children heard the same stories in similar-sized groups but afterwards did not engage in any of the linguistic activities. This study also obtained mixed results, with an improvement in false belief understanding being found only in the older, 4-year-old group but no improvement seen in the 3-year-olds. Taken together, the findings from the above-mentioned studies suggest that training children in groups on mental state language through reading stories and discussion with the teacher for several weeks does not have a robust effect on ToM understanding.

It is important to mention that while the majority of studies within the conversational training approach have focused on false belief and applied false belief tasks to measure effects, a few of them have included other tasks. For instance, Astington and Peskin (2004) also measured the children's ability to comprehend and produce cognitive verbs. For their part, Ornaghi et al. (2011) examined their child participants' abilities even more comprehensively: in addition to performing false belief and a mental state verb comprehension tasks, participants were tested for potential gains in affective aspects of social cognition and pragmatic competence. To evaluate the affective aspect of social cognition, a test measuring children's comprehension of the nature, causes, and regulation of emotions was used. To measure pragmatic competence, a brief 12-item pragmatic judgment test (Carrow-Woolfolk, 1999) was used in which the children were asked to make judgements about the appropriateness of responses to particular situations. Like for ToM, results for emotion understanding and pragmatic competence revealed different effects depending on age. On the one hand, training had a beneficial effect on the emotion understanding only in the younger, 3-year-old children. On the other, while there was a significant conversational training effect on pragmatic competence for 3-year-old children relative to an untrained control group, for the 4-year-olds both the experimental and control groups significantly improved their performance.

In sum, up thus far inconclusive findings have been obtained with regard to the effects of conversation-based training with mental state language on ToM. However, as mentioned above, most studies have focused on false belief measures and have not assessed other

[1] In the articles listed above, mental state terms are also referred to as *inner state terms*, *metacognitive language* and *mental lexicon*. The studies by Guajardo and Watson (2002) and by Astington and Peskin (2004) focus on the use of cognitive mental state terms (e.g., *know*, *think*), while the studies by Esteban et al. (2010) and Ornaghi et al. (2011) include both cognitive and emotional states.

complementary competencies such as emotion understanding and pragmatic abilities.

1.2. The need for a multidimensional approach for assessing training effects

As mentioned, though awareness of false belief has been the primary focus of attention in prior research on social cognition, authors such as Wellman (2018) have criticized the false belief paradigm as an overly categorical measure for the very profound ability to attribute mental states to another person. Given these issues, a multidimensional test battery would provide a more comprehensive understanding of a child's developing social cognition and pragmatic abilities. Thus, the current training study will adopt a multidimensional approach that incorporates not only false belief tests but also a mental state verb comprehension test, and other measures of social cognition and pragmatic abilities. This strategy will allow for a more fine-grained empirical comparison of the efficacy of experimental interventions across tasks.

Although pragmatic skill measures represent a comprehensive assessment of social ability, so far little attention has been devoted to assessing the potential positive effects on these skills of mental-state-related conversational interventions. As mentioned above, the study by Ornaghi et al. (2011) assessed gains in the pragmatic competence of 3- and 4-year-old preschoolers, measured through a short pragmatic judgement test. Other intervention studies on pragmatic abilities have focused on specific and fairly advanced receptive pragmatic skills rather than on expressive pragmatic competence, training children, for example, on nonliteral meanings such as irony (Pexman, Reggin, & Lee, 2019; Szücs & Babarczy, 2017), metaphors (Białecka-Pikul, 2010; Cortés, Cobos, & Tarbox, 2018), or inferencing (Davies, McGillion, Rowland, & Matthews, 2019). In addition, most prior research on pragmatic training has concentrated on intervention programs designed to improve the pragmatic abilities of children with disabilities, specifically children with Autism Spectrum Disorder (for a recent review see Parsons, Cordier, Munro, Joosten, & Speyer, 2017; see also Adams et al., 2012), and also adults after traumatic brain injury (Bosco, Parola et al., 2018; Gabbatore et al., 2015). Nevertheless, as noted in Bambini et al. (2016), the evaluation of pragmatic abilities is still rarely carried out in clinical settings.

Thus, no study so far has fully assessed the potential effects of a focused classroom-based training intervention on the expressive pragmatic abilities of typically developing children using a comprehensive pragmatic test battery. As for mental state verb comprehension and production, as we have seen, some previous studies have assessed the effect of conversational training on mental state verb comprehension along with other ToM measures (Astington & Peskin, 2004; Ornaghi et al., 2011). By the same token, other studies adopting a conversation-based approach have assessed affective social cognition. For example, in Tenenbaum, Alfieri, Brooks, and Dunne (2008)) 5- to 8-year-old children were individually trained with story reading followed by explanatory conversations for four weeks, yielding an increase in emotion understanding. Similarly, a study by Grazzani and Ornaghi (2011) found that training 3- to 5-year-old preschool children in groups with conversational language games for two months had a significant effect on their emotion comprehension. However, the present study is to our knowledge the first to assess the impact of conversational training on all three of these social cognition factors as well as pragmatic competence.

1.3. Aims of the present study

The first overarching aim of the present study is to assess the potential positive effect of a four-week group-based conversational classroom intervention not only on false belief understanding, but also on two more aspects of social cognition (mental state term comprehension and emotion understanding) as well as expressive pragmatic abilities. We hypothesize that conversational mental state training for preschoolers based on talking about mental and emotional states will enhance both their social cognition and their pragmatic skills. As mentioned above, we believe that such a multidimensional approach will provide richer insights into the effectiveness of training than a unidimensional approach based exclusively on false belief tasks.

The second aim of the study will be to assess the potential effects of two different paradigms for conversational mental state classroom training, one involving conversation and verbal expression alone (*non-embodied* training) and the other involving both verbal and behavioral expression (*embodied* training), where children are asked to act out mental state concepts using gesture and prosody (see Pexman, 2019 for the role of embodiment in conceptual development). To our knowledge, this will be the first study to assess the role of embodied conversational training. Given the strong relationship between pragmatic development and the development of multimodal communication, it is important to ask whether embodied forms of training are also effective for boosting pragmatic abilities.

A growing body of research has highlighted that pragmatic development in children goes hand-in-hand with their use and understanding of gesture and prosody. First, many studies have investigated the co-development of gesture, prosody, and pragmatics at early ages (see Esteve-Gibert & Prieto, 2018; Goldin-Meadow, 1998, 2014, for reviews), suggesting that in order to communicate their intentions preverbal children rely heavily on multimodal aspects of communication such as prosody (Carpenter, Akhtar, & Tomasello, 1998; Hoicka & Wang, 2011; Sakkalou & Gattis, 2012), facial expression (Oller et al., 2013; Scheiner, Hammerschmidt, Jürgens, & Zwirner, 2002), eye gaze (Behne, Carpenter, & Tomasello, 2005; Farroni, Csibra, Simion, & Johnson, 2002), manual gestures (E. Bates, Camaioni, & Volterra, 1975; Camaioni, Perucchini, Bellagamba, & Colonnesi, 2004; Tomasello, Carpenter, & Liszkowski, 2007), and their combination (Aureli et al., 2017; Esteve-Gibert, Prieto, & Liszkowski, 2017; Grünloh & Liszkowski, 2015). Second, several studies have shown that prosody and gesture play an important precursor role not only at early stages of language development but also later on, during the preschool years, when children undergo important changes in their pragmatic development (see Hübscher & Prieto, 2019, for a review). Research has provided empirical evidence that multimodal cues pave the way for the acquisition of a variety of pragmatic skills during the preschool years and beyond, such as speech act marking (Rusiewicz & Esteve-Gibert, 2018), focus and informational structure (Esteve-Gibert, Lœvenbruck, Dohen, & D'Imperio, 2017; Frota & Butler, 2018), epistemic stances (Armstrong,

Esteve-Gibert, Hübscher, Igualada, & Prieto, 2018; Bartz, 2017; Hübscher, Esteve-Gibert, Igualada, & Prieto, 2017; Hübscher, Vincze, & Prieto, 2019; Kim, Paulus, Sodian, & Proust, 2016), emotional states (Berman, Chambers, & Graham, 2016; Berman, Graham, Callaway, & Chambers, 2013; Berman, Graham, & Chambers, 2013; Berman, Chambers, & Graham, 2010; Khu, Chambers, & Graham, 2018), politeness (Hübscher, Prieto, & Wagner, 2020; Hübscher, Garufi, & Prieto, 2019), pragmatic inferences and irony (González-Fuente, 2017; Laval & Bert-Erboul, 2005), and discourse and interactional abilities like turn-taking (Casillas & Frank, 2012) and narratives (Colletta et al., 2015; Goldin-Meadow & Wagner, 2005; Graziano, 2014; Llanes-Coromina, Vilà-Giménez, Kushch, Borràs-Comes, & Prieto, 2018; McNeil, Alibali, & Evans, 2000; Sekine & Kita, 2015).

Despite this well-established relationship between pragmatics and multimodal development, however, little is known about whether multimodal training interventions based on face-to-face conversations can contribute to enhancing social cognition and pragmatic skills. Since children benefit from multimodal cues and are sensitive to prosodic and gestural markers, we would expect that enacted training interventions that encourage the production of embodied, multimodal prosodic and gestural expressions of mental states and emotions will enhance their acquisition of social cognition and pragmatic skills. This study is intended to shed light on this issue. Though previous studies have assessed conversational mental state training interventions (with mixed results), no previous study has assessed the potential of embodied training interventions, that is, conversation coupled with gesture and prosody. If comparable training effects can be achieved with the embodied training paradigm we apply in the present study, this would be of particular relevance regarding children whose verbal language skills are weak, whether due to their developmental stage, linguistic impairments or a different linguistic background.

2. Methods

The experiment consisted of a four-week between-subjects classroom-based training paradigm with a pretest and posttest design. Children 3 and 4 years of age were assigned to one of three training conditions: non-embodied conversational, embodied conversational, and control. From now on we will use the shorter term non-embodied and embodied conditions. This age range was selected because a large body of research has demonstrated that critical development of social cognition occurs during the early preschool period (e.g., Wellman, 2018).

2.1. Participants

Participants in the study were 3- to- 4-year-old Catalan-speaking children (*M* age = 44.75 months, *SD* = 3.27). All participants were preschoolers from two Catalan public schools located in the middle-income district of Sant Martí in the central district of Barcelona, where the population is largely Catalan-Spanish bilingual.[2] The main language of instruction in these schools is Catalan. Our training scheme was embedded within the school curriculum, that is, it was carried out during school hours, and a great majority of families (94 %) provided written consent for their children's participation. Parents of participating children also completed an occupational status questionnaire, which yielded a mean ISEI score of 60.90 (*SD* = 13.23; see Ganzeboom, De Graaf, Treiman, & De Leeuw, 1992), confirming overall middle class status, as well as a language questionnaire regarding the daily exposure of their child to Catalan (mean overall exposure time = 59 %, *SD* = 20.83). All children were typically developing children and had no history of speech, language, or hearing difficulties.

Although initially 117 students were enrolled, 34 ultimately had to be excluded from analysis, for different reasons. First, 15 of the initial recruits failed to meet the 20 % minimum score in a screening test of Catalan vocabulary proficiency (see section 2.3.1 below). Subsequently, two children missed school and dropped out of the experiment, while 17 others were too restless or inattentive to participate during either the pretests or posttests. This left a final study sample of 83 children, of whom 46 were female and 37 male.

2.2. Materials

This section will describe first the materials used in the pretest and posttest sessions, then the training materials.

2.2.1. Pretest and posttest tasks

The following four abilities were assessed at pre- and posttest: (1) pragmatic competence; (2) false belief understanding; (3) mental state verb comprehension; and (4) emotion understanding. The materials used to measure these abilities are described in detail below.

2.2.1.1. Pragmatic competence. Children's expressive pragmatic competence was assessed by means of the Audiovisual Pragmatic Test (APT) (Pronina, Hübscher, Vilà-Giménez, & Prieto, 2019), which was designed to assess pragmatic abilities in Catalan-speaking children from early childhood to adolescence. The test was developed to assess expressive pragmatic abilities in terms of the appropriateness of a spoken reaction to a discourse prompt. The general design, item coverage, and elicitation procedure of the APT is based on currently used pragmatic tests for children (e.g., TOPL-2 (Phelps-Terasaki & Phelps-Gunn, 2007) and CASL–2 (Carrow-Woolfolk, 2017)). In broad terms, the procedure is based on the Discourse Completion Task method, in which an everyday social context is

[2] According to the Barcelona City Council's Department of Statistics database for 2018, which can be accessed at https://www.bcn.cat/estadistica/catala/dades/anuari/cap06/C0616010.htm, 96% of the population of Sant Martí understands Catalan and 79% speaks it.

described and then the participant is asked to respond with the speech act that would be most natural in that context. The APT contains a total of 47 items, accompanied by illustrations, that represent different contexts that might plausibly occur in a child's everyday life (see Fig. 1 for an example).

For the purposes of this study, only the first 35 items were selected, given that the test takers were 3- to 4-year-old children who would be unlikely to have encountered in their daily lives the more complex pragmatic situations presented in the last 12 items. The items were presented in a fixed order.

2.2.1.2. False belief understanding. The first type of measure used to assess children's social cognition tested their false belief understanding. Two tests were used for this purpose. The first one was the classic "Sally and Anne" change-of-location task (Baron-Cohen et al., 1985), in this instance adapted to Catalan (Armstrong et al., 2018). The task was presented in video format, with the video showing short sequences involving puppets (see Fig. 2). The child was asked two questions, the first a false belief test question ("Where will the princess look for the ball?"), the second a control question ("Where is the ball really?").

The second measure was the "Smarties" unexpected content task (Gopnik & Astington, 1988). In this task, the experimenter gives the child the characteristic tube in which Smarties chocolate candies are sold. The child opens the tube and finds that it contains pencils, not candies. The experimenter then asks the child two questions, in this order: a self-test question ("What did you think was in the box before you opened it?"), an other-person-test question ("What would your friend think was inside the box before it was opened?"). In this case, the Smarties tube used in the original task was replaced with a similar tube used by a local candy brand, *Lacasitos*, which would be familiar to Catalan children.

2.2.1.3. Mental state verb comprehension. A second measure of the cognitive aspect of social cognition, the conceptual comprehension of cognitive mental states, was assessed using the Metacognitive Vocabulary Test (MVT) (Pelletier & Astington, 1998), which is designed for 3- to 7-year-old children. The test consists of a total of 12 short stories accompanied by images and measures children's ability to discriminate among mental state verbs describing the main character's state of mind (see Fig. 3). Following Ornaghi et al. (2011), only the first six stories were used due to the young age of the participants. The mental state verbs assessed in the first six stories include *know*, *remember*, *guess*, *figure out*, *wonder*, and *forget*. For the purposes of this study, the test was translated from English into Catalan by the second author of the study, with the translation checked by a Catalan native speaker.

2.2.1.4. Emotion understanding. To assess the affective aspect of social cognition, the Emotion Matching Task (EMT), originally designed for English-speaking preschool-aged children by Izard, Haskins, Schultz, Trentacosta, and King (2003) was used. The EMT has been translated and adapted for preschool children in Spain (Alonso-Alberca, Vergara, Fernández-Berrocal, Johnson, & Izard, 2012), and the Spanish version has also been translated into Catalan by the same authors. The EMT measures the emotion comprehension of 3- to 6-year-olds and focuses on four basic emotions: happiness, sadness, anger, and fear/surprise. In the present study, only the two most discriminating parts of the EMT were administered: the expression-situation matching subtest and the expression labeling subtest. The expression-situation matching subtest consists of a total of 12 items. The child is asked to indicate which of the four pictures expressing diverse emotions or a 'neutral' face match a described situation (see Fig. 4, left panel). The expression labeling subtest consisted of a total of 12 pictures that assess the child's ability to name emotions (see Fig. 4, right panel).

2.2.2. Training session materials

The materials for the eight training sessions were taken from the *The Adventures of Jack and Theo* (Ornaghi & Grazzani, 2020), a set of stories specifically designed for and used in the training study by Ornaghi et al. (2011). The set comprises 16 stories arranged in order of increasing difficulty that narrate the adventures of two characters, Jack and Theo. The first eight stories were translated into Catalan for use here (a sample is provided in the Appendix A). Then high-quality audio recordings were made of the stories being read aloud in child-directed speech by a preschool teacher under the supervision by the second author to ensure that the engagement level expressed by the reader was uniform. Finally, eight short videos were created using the audio tracks accompanied by illustrations from the original materials by Ornaghi et al. (2011)[3].

The eight stories used in this study are enriched with mental state terms. A total of 12 mental state terms were trained: eight of them cognitive state terms and four of them emotional state terms (see Table 1). Each of the eight stories trained two terms, either one cognitive state and one emotional state, or two cognitive states. For example, in the sample story provided in the Appendix A, the two terms are *believing* (cognitive) and *getting upset* (emotional). Thus, the eight cognitive state terms each appeared once in the full set of eight stories, whereas the four emotional state terms each appeared twice.

Once the audiovisual prompt materials were ready, scripts were prepared for the trainer who would conduct the training sessions in the two experimental groups. In contrast with Ornaghi et al. (2011), the scripts for the training procedure were highly controlled. After listening to the story, the children in the experimental groups participated in a group interaction with the teacher. The key difference between the two experimental conditions consisted in the instructions given by the experimenter. In the non-embodied condition, children were asked to tell what the story character or they themselves say when they have a certain mental state, whereas in the embodied condition, children were (1) asked to tell the same but, importantly, (2) the experimenter also encouraged children *to show*

[3] All training materials including stories in video format in Catalan, scripts, instructions and examples of training sessions can be found on the specifically designed educational website https://entrenemlesemocions.wordpress.com.

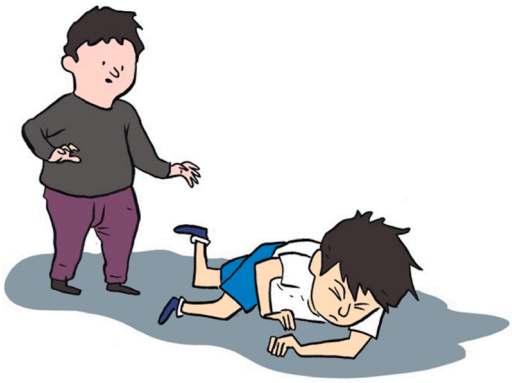


Fig. 1. Sample item from the Audiovisual Pragmatic Test showing text in English translation and illustration intended to elicit an expression of concern for a friend.

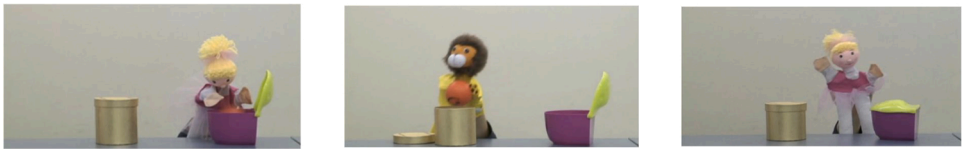


Fig. 2. Stills of a video sequence from the false belief task. From left to right: (1) the princess puts a ball in the purple box and leaves; (2) the lion moves the ball from the purple box to the gold box; (3) the princess comes back (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article).

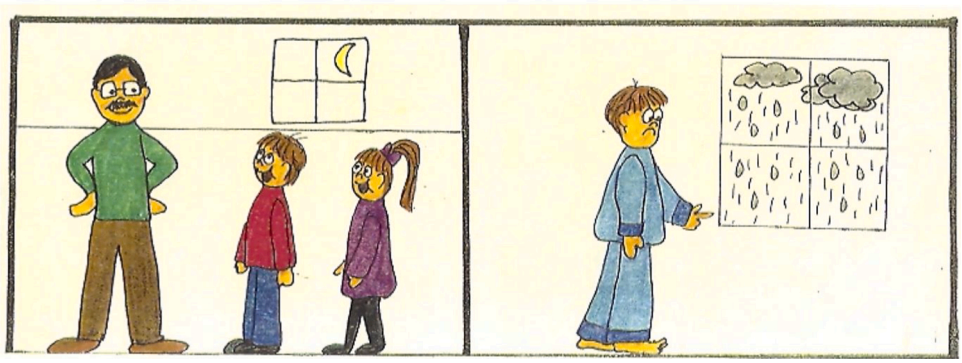


Fig. 3. Sample item from English version of the Metacognitive Vocabulary Test showing the illustration (left) and corresponding text to be read to the child (right), including the prompt question in which the child is asked to decide between two mental state verbs (in italics).

what the story character or they themselves do when they have this mental state. In this manner, the non-embodied group answered only verbally during the training, while the embodied group answered verbally and also enacted the answers. Since the training in both groups was kept the same amount of time, the embodied group necessarily practiced less in answering verbally, i.e., using mental state terms and focused primarily on the enactment of the different cognitive and emotional states.

As exemplified in Table 2, which contains part of the script for Story 8 (see Appendix A), all scripts followed the same general scheme. First, the trainer directs the children's attention to the first target term; once discussed, the second target term is introduced. Questions for both target terms also follow the same design: the children are asked a series of six questions in a fixed order. However, as can be seen in the table, the non-embodied training vs. embodied training variants of the script differ with regard to the second and fifth of these questions: questions in the non-embodied variant merely ask the children to report what story characters said, whereas questions in the embodied variant ask the children to report what characters said and to show what they did.

The two variant scripts were previously memorized by the trainer, a professional primary school teacher and stage actress. She was also instructed to provide positive feedback for every appropriate response, which would be either verbal or multimodal depending on whether she was working with the non-embodied or embodied group. Also, if it was apparent that the children were having troubling responding, the trainer was instructed to ask additional questions to encourage participation. In all instances, the children were to be given a model answer regardless of how they performed. For example, for the term *getting upset* trained in Story 8, the trainer would

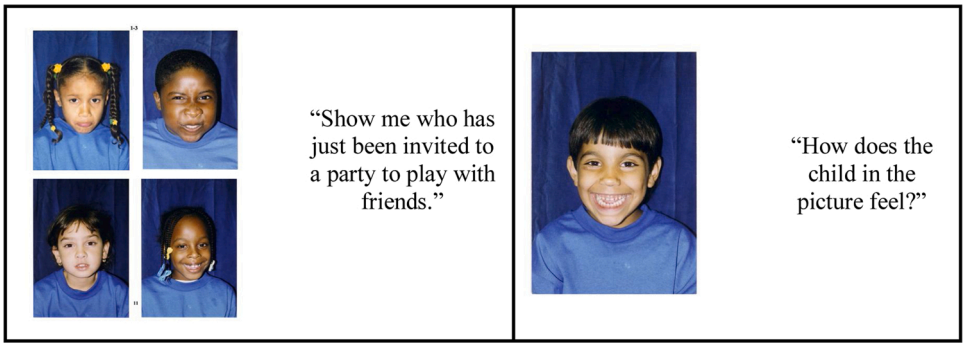


Fig. 4. Sample items from the English version of the Emotion Matching Task showing texts to be read to the child and accompanying illustrations. Left panel shows a sample from the expression-situation matching subtest, with children displaying, clockwise from top left, sad, angry, happy, and neutral expressions. Right panel shows a sample from expression labeling subtest.

Table 1
Target cognitive and emotional state terms contained in the training materials.

Cognitive state terms	Emotional state terms
wanting	*getting delighted*
remembering	*getting upset*
knowing	*getting angry*
thinking	*getting scared/surprised*
believing	
deciding	
doubting	
wondering	

Table 2
English translation of partial script for training session corresponding to Story 8, showing the variants corresponding to the two experimental conditions. Target emotional state term: *getting upset*.

Non-embodied condition	Embodied condition
1. Do you remember that in the beginning of the story Theo asked Sara the Sea-Turtle: "What happened? Why are you crying?" Today we are going to play using the words "getting upset".	
Do you remember why Sara was upset and why she was crying?	
2. What did Sara say when Theo found her stuck in a hole?	2. What did Sara do when Theo found her stuck in a hole? Can you tell me what she said?
3. What makes you upset?	
4. What makes your friends upset?	
5. What would your friend say if his favorite toy broke?	5. What would your friend say if his favorite toy broke? What would he do in this situation? Can you show me?
6. Well, children, you've been very good. We have been playing using the words "___". (Children finish the sentence)	

give an example of the verbal answer "I'm so upset, I got stuck in a hole and I can't get out" in the non-embodied condition. In the embodied condition, the experimenter would provide the same verbal answer but also use appropriate prosodic and gestural cues to express being upset as she spoke; the gestural and prosodic patterns were stipulated for each of the target mental states and rehearsed by the trainer prior to the experiment (see Fig. 5). Further, in the embodied condition, children were to be encouraged at certain points to enact the target mental state as a group.

The materials and training sessions for both conditions were previously piloted on a group of 19 3-year-old children, who did not take part in the subsequent main study.

2.3. Experimental procedure

A diagram of the experimental procedure can be seen in Fig. 6.

2.3.1. Screening measure

Since Catalan was to be the language employed during the training intervention, participants were given a screening test to confirm that they all had a sufficient command of Catalan to successfully understand the tasks. The measure used was an expressive one-word

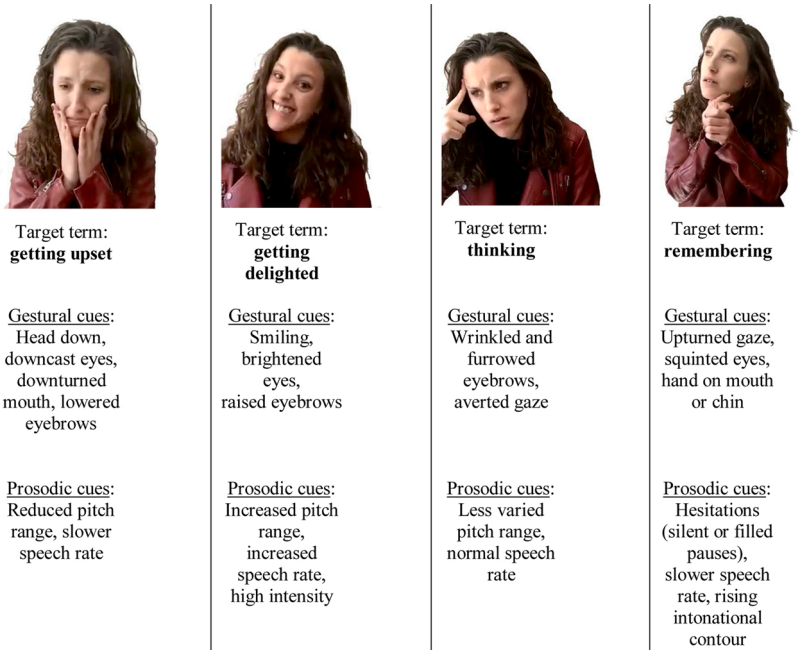


Fig. 5. Examples of the trainer modeling target terms and the prosodic and gestural cues employed for each term.

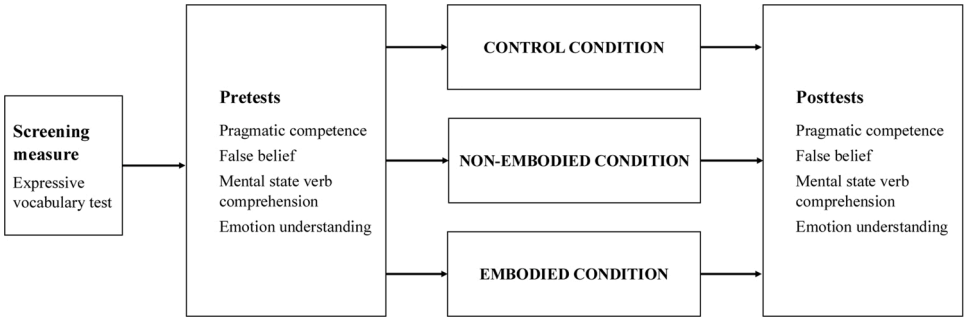


Fig. 6. Experimental procedure.

vocabulary test designed to measure the lexical knowledge of Catalan-speaking children aged 0–6 by means of picture-naming tasks (Saborit Mallol, Julián Marzá, & Navarro Lizandra, 2005). The stimuli consist of 30 pictures of common objects such as a tree or a coat. Each participant is given credit for every correctly named item, and the total score (from 0 to 30) is then normalized on a 0–100 scale. For purposes of the present study, a minimum score of 20 % was set as the eligibility threshold for participation. As noted above in Section 2.1, data from 15 children were excluded from analysis on this basis.

2.3.2. Pretest and posttest procedure

Pretest and posttest, which took place respectively before and after the four-week training intervention, followed identical procedures. Each child was tested individually at their corresponding school during school hours. The child was brought from his/her classroom to a quiet room where he/she was seated at a table beside an experimenter, either the first author or one of three trained research assistants. All testing was conducted in Catalan and the full session lasted around 30–35 min. Testing followed the same order for all children, with emotion understanding being tested first, followed by mental state verb comprehension, false-belief understanding, and pragmatic competence.

2.3.2.1. Pragmatic competence. The child faced the computer screen where the images for each item in the APT were presented. Prior to the APT proper, two familiarization trials were carried out. For each item, the examiner described the social situation represented in the picture in a lively fashion appropriate for a preschooler while the child looked at the illustration displayed on the computer screen. The examiner then asked the child to respond appropriately as if he/she was taking part in the situation. The total duration of the APT was around 15 min.

Pragmatic abilities were assessed in terms of the speech act appropriateness of the answer provided by each child, from an illocutionary and social point of view. Each answer was given a score ranging from 0 to 2. A score of 2 was given if the answer was a target speech act of high pragmatic quality, meaning that the child managed to utter the intended speech act and showed social adjustment. For example, in a situation in which the child was prompted to ask for a piece of cake, if the child uttered a sentence along the lines of "Can I have a piece?" this was given a score of 2, since it was the target requesting speech act and the child had reacted appropriately from a social point of view. By contrast, a score of 1 was given if the child managed to produce the target speech act but the answer was not quite socially appropriate. Depending on the situation, this might mean that the child said too much or too little or answered too directly. For example, if when asking for a piece of cake the child said something like "Give me a piece", the answer was given a score of 1 since the child was able to express a request but was too demanding or insufficiently polite in doing so. A minimum score of 0 was given if the child did not provide any answer, provided an unrelated answer, or answered in a socially inappropriate way. For example, if when refusing the cake the child said, "Give me more", the answer was scored as 0 since the child had clearly not understood the situation. If when prompted to express concern for a friend who had just tripped and fallen down the child did not manage to express worry and simply said something like "You fell" in a blunt way, a score of 0 was given. On the other hand, an appropriate answer to this item like "Are you alright? Do you need help?" was scored as 2.

Finally, each participant's scores were averaged to produce a general pragmatic appropriateness score, the maximum score possible being 70 (35 items × maximum 2 points).

2.3.2.2. False-belief understanding. In the first task, the child was shown a video and then asked false-belief and control test questions; a score of 1 was given for each correct answer to the two questions. In the second task, the child was shown a *Lacasitos* tube and then asked self-test and other-person-test questions; a score of 1 was given for each correct answer to the two questions. The administration of the two tasks took approximately 5 min in total. Scores from the two tasks were added to yield a composite ToM score ranging from 0 to 4

2.3.2.3. Mental state verb comprehension. First, two training items were administered so that the child would understand the test procedure. Then the six stories from the MVT were read to the children. After each story they were asked to select which of two cognitive verbs correctly described the main character's state of mind. The duration of the test was around 5 min. One point was awarded for each correct verb choice. The range of possible total scores was therefore from 0 to 6.

2.3.2.4. Emotion understanding. In the 12 trials of the expression-situation matching part of the EMT, the child was shown four pictures of children expressing diverse emotions and was then asked to indicate which picture matched a described situation. In the expression labeling part of the EMT, after being shown a picture, the child was required to identify the emotion depicted. Altogether the EMT took approximately 10 min. Following standard procedure, children were awarded one point for each correct answer in the first part of the test, yielding a possible maximum score of 12 points, and were awarded one or two points, depending upon accuracy in labeling, for each correct answer in the second part, yielding a possible maximum score of 24 points. The range of possible total scores was thus from 0 to 36.

2.3.3. Training procedure

Following the administration of the pretest, the participating children were assigned to one of three different conditions. In order to guarantee a homogeneous distribution of children across the three conditions, a special algorithm[4] was written to control for potential differences in the abilities measured. Once the children were distributed into groups according to the results of the algorithm, a separate one-way ANOVA analysis was run for each pretest measure, results on the expressive vocabulary screening test, and age to check that the three groups were not statistically different. For comparison purposes all test scores were normalized to a 0–100 scale. As can be seen in Table 3, ANOVA analysis confirmed that the groups did not significantly differ ($p > .05$ in all cases).

2.3.3.1. Intervention sessions. Children underwent one of three possible interventions in small groups of about 12 children over four weeks in eight sessions lasting between 20 and 25 min. Sessions occurred twice a week on nonconsecutive days. They took place in a quiet classroom at the children's school and were videotaped using three cameras simultaneously. One camera was directed at the experimenter, another at the children, and the third provided an image of the full setting.

The three between-subject conditions distinguishing the groups were: 1) no specific training in social cognition or pragmatic competence skills (the control group); training consisting of experimenter-led conversations which encouraged children to reflect on the mental states of themselves and others (the non-embodied experimental group); and 3) training consisting of conversations which encouraged children to not only reflect on but also act out mental states (the embodied experimental group). In all three conditions, at the beginning of each session, the children watched a video telling a story enriched with mental state vocabulary (see Table 1). Children in the control group then simply carried out non-conversational activities related to the story, like drawing pictures or solving puzzles depicting sea creatures and other animals, for the remainder of the session. By contrast, children in the experimental groups received one or the other variant of the conversational training intended to enhance their social cognition and pragmatic competence

[4] The python algorithm that was used to assign participants to experimental conditions while taking into account children's class distribution and preserving general similarity of conditions across groups is freely available on Github repository.

Table 3
P values resulting from three-group ANOVA comparison of pretest measures, screening test scores, and mean age.

Measures	*p*
PRETEST:	.72
Pragmatic competence	.72
False-belief understanding	.62
Mental state verb comprehension	.46
Emotion understanding	.98
SCREEING TEST	.53
MEAN AGE	.56

skills. Unlike the control group, these two groups followed scripted procedures. However, the time taken up by the intervention sessions was kept strictly uniform across all three groups.

We will now describe the procedures followed in the two experimental groups. The same trainer performed the training for both groups. The first author also remained present during each session in order to monitor the time and ensure proper execution of the script by the trainer. Depending on the group, the trainer followed one or the other variant of the script. In the non-embodied training group, this meant involving the children in a question-driven conversation in which they were asked to reflect on a particular mental state in not only themselves but also other people. In the embodied group, children were asked to both reflect on the target concept and also act it out using prosody and gesture. In each session, this procedure was followed for two separate mental states, linked through one of the stories involving Theo the Shark. The two photos in Fig. 7 show the experimental setup of training sessions in the non-embodied and embodied conditions.

Throughout each training session, the trainer encouraged participants to use the target terms as much as possible. She also motivated children to participate actively in the conversation and ensured that all of them were involved in the activity.

2.4. Statistical analyses

The data from pretest and posttest were compared within and across the three groups using a linear mixed-effects model in R version 3.6.1 (R Core Team, 2019) with the *lme4* package (Bates, Mächler, Bolker, & Walker, 2015). A total of four linear mixed-effects models were run with each of the tests measuring different abilities assessed (pragmatic competence, false belief understanding, mental state verb comprehension, and emotion understanding) as the dependent variable. Condition (three levels: control condition vs. non-embodied training condition vs. embodied training condition), Test (two levels: pretest vs. posttest), and their interaction were set as fixed factors. Predictors were coded using mean-centered effects coding. Likelihood ratio tests suggested that inclusion of the Age variable significantly improved the fit of all four models, therefore Age was also set as a fixed factor. A random intercept was set for participants; random slopes were not included due to non-convergence of the model. The post-hoc pairwise comparisons were computed with the *emmeans* package (Lenth, Singmann, Love, Buerkner, & Herve, 2018). Partial eta-squared values for effect sizes were extracted using the *sjstats* package (Lüdecke, 2016). Using likelihood ratio tests, the gender variable was found to be not statistically significant for models and was left out of future subsequent analyses.

The analyzed dataset is available on Github repository (https://github.com/mariia-pronina/boosting_pragmatic_abilities).

3. Results

3.1. Pragmatic competence

For pragmatic competence, the linear mixed-effects model analysis revealed a significant main effect of Test ($p < .001$) and Age ($p < .001$). There was no significant main effect of Condition. The full estimates are given in Table 4. The estimated mean percentages of correct responses were 21.15 ($SD = 15.06$) for the pretest and 25.66 ($SD = 13.84$) for the posttest. Further, a significant interaction between Condition and Test was found for pragmatic competence ($p = .015$), showing a difference in the pragmatic scores among

Fig. 7. Training session photographs. The left panel shows the non-embodied condition, and the right panel the embodied condition.

control, non-embodied, and embodied conditions depending on the time of the administration of the test (pretest vs. posttest). Post hoc pairwise comparisons likewise showed a difference in the pragmatic scores among control, non-embodied, and embodied conditions depending on the time of the administration of the test. Specifically, post hoc pairwise comparisons revealed that the difference between pretest and posttests pragmatic scores was statistically significant in the two experimental groups ($p = .005$ for non-embodied group, $p < .001$ for embodied group), with better scores in the posttest, but not in the control group ($p = .666$).

To illustrate this, the mean observed percentages of correct responses by test time for each training condition are given in Table 5, along with estimates and partial eta-squared values. These patterns are graphically illustrated in Fig. 8.

3.2. False belief understanding

The linear mixed-effects analysis showed no significant main effect of Condition, Test, or Condition × Test interaction for false belief understanding; only Age turned out to be significant ($p = .001$). The full estimates are given in Table 6.

Additional analyses that took into account the type of false belief task and initial ToM performance at pretest did not lead to substantively different results.

3.3. Mental state verb comprehension

With regard to mental state verb comprehension, no significant differences between pretest and posttest scores were found for any of the groups. There was no significant main effect of Test or Condition and no significant Condition × Test interaction; Age was the only significant predictor ($p = .024$). The full estimates are given in Table 7.

3.4. Emotion understanding

The linear mixed-effects analysis revealed a significant main effect of Test ($p < .001$) and a marginal effect of Age ($p = .051$) for emotion understanding. No significant main effect of Condition was found, nor any significant interaction between Condition and Test. The full estimates are given in Table 8.

Mean percentages of correct responses on emotion understanding pretests and posttests as well as estimates and partial eta-squared values for each condition are given in Table 9.

4. Discussion and conclusions

The results obtained in the present study reveal that preschool children who received training in either of the two experimental conditions (non-embodied or embodied conversational training) significantly improved their pragmatic skills as compared to a control group that received no training. Pairwise comparisons confirmed that the pragmatic competence posttest score improved relative to from the pretest score only in the two experimental conditions. On the other hand, with regard to social cognition skills related to false belief understanding, mental state verb comprehension, and emotion understanding, these remained unchanged from pretest to posttest in all conditions.

One of the clear findings of the study is thus that training preschool children through either non-embodied or embodied conversation about mental state leads to an improvement in their pragmatic score, suggesting that children benefit from a short classroom training in acquiring pragmatic competence. It is worth noting that the duration of training sessions was kept constant across the two different experimental groups (and also the control group). As a consequence, while the children in the non-embodied group had solely practiced the use of mental state terms, the children in the embodied group had less opportunity to practice using these verbs since part of the session was devoted to acting them out. Crucially, however, the results show that the two types of training were equally effective. The fact that children in the embodied condition did as well as (and even slightly better than) the children in the non-embodied group, despite less exposure to the verbal training, is a significant finding, because it means that bodily actions are as effective as words and thus embodied training has the same training power as verbal training. This finding has several important implications. It suggests, for example, that embodied training may be successfully used for mental state training in children with low verbal abilities. Moreover, although pairwise comparisons revealed no significant differences between the non-embodied and embodied groups in the posttest, a larger effect size was observed in the embodied group. This suggests that embodied training interventions may yield even better results in future studies if the exposure of children to both embodied training and mental state talk is more prolonged and intensive. Thus, our findings support a multimodal view of language highlighting the role of embodiment in training mental states.

Table 4
Model specification and estimates for pragmatic competence.

Fixed effects	β	*SE*	*t*	*p*
Intercept	23.41	1.38	16.92	< .001
Condition	−0.51	1.40	−0.362	.718
Test	2.26	0.56	4.024	< .001
Age	5.17	1.40	3.702	< .001
Condition: Test	1.40	0.56	2.484	.015

Table 5
Mean (SD) percentages of correct responses on pretest and posttest, β, *SE*, *t*, and *p* values and effect sizes (ηp^2) for the test of pragmatic competence, broken down by condition.

	Pretest	Posttest	β	*SE*	*t*	*p*	ηp^2
Control condition	22.76 (16.98)	23.57 (16.41)	−7.53	1.87	−0.433	.666	.009
Non-embodied condition	20.95 (13.83)	26.67 (11.10)	−5.71	1.97	−2.897	.005	.202
Embodied condition	19.51 (14.30)	27.03 (13.41)	−0.81	2.01	3.746	< .001	.256

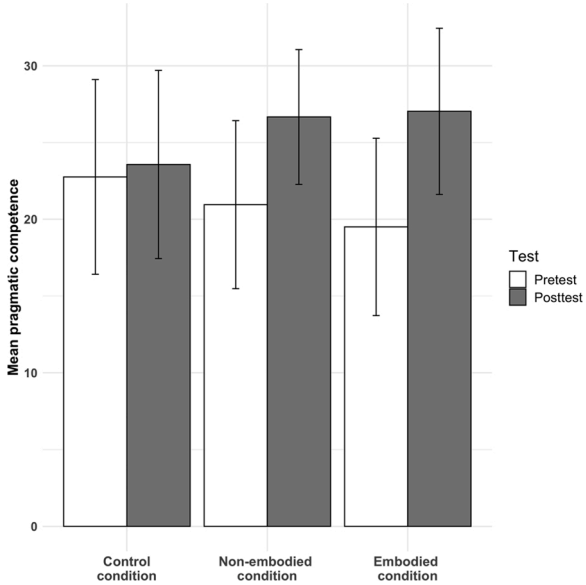


Fig. 8. Mean overall pragmatic score, broken down by condition (control vs. non-embodied vs. embodied) and test (pretest vs. posttest). Error bars represent 95 % confidence intervals of the means.

Table 6
Model specification and estimates for false belief understanding.

Fixed effects	β	*SE*	*t*	*p*
Intercept	39.71	1.90	20.893	.000
Condition	1.17	1.92	0.607	.546
Test	0.55	1.51	0.367	.715
Age	6.88	1.91	3.602	.001
Condition: Test	−0.03	1.52	−0.018	.985

Table 7
Model specification and estimates for mental state verb comprehension.

Fixed effects	β	*SE*	*t*	*p*
Intercept	50.11	1.71	29.389	.000
Condition	−1.17	1.72	−0.681	.498
Test	0.11	1.51	0.098	.922
Age	3.96	1.72	2.308	.024
Condition: Test	−1.24	1.14	−0.089	.280

While the two types of conversational training had beneficial effects on pragmatic abilities, the performance on the social cognition tasks surprisingly did not differ across conditions in the posttest. With respect to false belief understanding, none of the three groups (control, non-embodied, and embodied) showed any significant increase in false belief performance. Modified analyses that included only children who performed achieved at most a 50 % score in the false belief tasks in the pretest or took into account the type of false belief task did not improve the statistical model. Our results contrast with the positive results of other more focused training interventions that essentially provided children with feedback about the accuracy of their responses (Appleton & Reddy, 1996; Lecce, Bianco et al., 2014; Melot & Angeard, 2003; Slaughter, 1998; Slaughter & Gopnik, 1996). It also contrasts with other types of training

Table 8
Model specification and estimates for emotion understanding.

Fixed effects	*β*	*SE*	*t*	*p*
Intercept	5577	1.45	38.431	< .001
Condition	1.15	1.46	0.787	.433
Test	8.11	1.03	7.885	< .001
Age	2.90	1.46	1.980	.051
Condition: Test	1.52	1.03	1.478	.143

Table 9
Mean (SD) percentages of correct responses on pretest and posttest, *β, SE, t*, and *p* values and effect sizes (*ηp2*) on the test of emotion understanding, broken down by condition.

	Pretest	Posttest	*β*	*SE*	*t*	*p*	*ηp2*
Control condition	47.50 (16.69)	61.20 (17.12)	−13.7	3.42	−4.010	< .001	.303
Non-embodied condition	48.25 (18.15)	62.24 (14.64)	−14.0	3.60	−3.884	< .001	.214
Embodied condition	47.33 (17.59)	68.59 (13.96)	−21.3	3.67	−5.791	< .001	.539

on specific syntactic structures (Hale & Tager-Flusberg, 2003; Lohmann & Tomasello, 2003). However, in line with some conversation-based training studies which reported mixed or negative results (Astington & Peskin, 2004; Esteban et al., 2010; Guajardo & Watson, 2002; Ornaghi et al., 2011), our training study showed no extra learning effects in either of the two experimental conditions. This negative finding might be explained by the nature of false belief tasks, which are rather short and use an all-or-nothing evaluation approach, whereas children acquire an understanding of mental states gradually (see Wellman & Liu, 2004).

Concerning mental state verb comprehension, as in the case of false belief understanding, the results showed that training did not lead to improved performance in any of the groups, whether experimental or control. This is consistent with previous studies reporting negative results (Astington & Peskin, 2004). The present findings confirm that 3- to 4-year-olds still cannot fully discriminate between different mental state verbs, as the percentage of correct choices found here was at chance level (around 50 %) both at pretest and posttest. This suggests that perhaps this test was simply too difficult for young preschoolers (for a similar line of argumentation in explaining children's failure on the mental state verb comprehension test, see also Grazzani & Ornaghi, 2011).

Regarding the emotion understanding test, children in all three conditions performed significantly better on the posttest than on the pretest. Again, although all three groups improved, the differences in posttest scores across conditions were not significant. However, a larger effect size was observed for the embodied training group, indicating that the embodied group improved in a stronger fashion than the non-embodied and control groups and suggesting a trend for the embodied group to improve more. In our view, the fact that the control group also showed improvement can be explained by the nature of the specially designed stories read to the children in all three conditions, in which emotional concepts played a prominent role. The results suggest that exposure to stories enriched with emotional state terms, even when the terms are not specifically trained and discussed with children, can enhance children's emotion comprehension. The same effect did not hold for mental state verbs since, as mentioned above, the test of mental state verb comprehension may have been too complicated for the preschoolers. Finally, an alternative explanation for the improvement in emotion understanding in all three conditions could be the learning effect obtained from the repetition of the task in the posttest phase.

From a methodological perspective, the current findings reveal the importance of adopting a multidimensional perspective in the assessment of social abilities in preschool children, and importantly include more comprehensive measures of the children's pragmatic competence, especially their expressive pragmatic abilities. To our knowledge, Ornaghi et al. (2011) is the only prior study to have assessed the effects of conversational training in promoting typically developing preschooler's pragmatic abilities (albeit receptive, not expressive, abilities), obtaining mixed results and different effects for different age groups. Importantly, instead of the short pragmatic judgement test used in Ornaghi et al.'s study, the present study used a comprehensive expressive pragmatic task in which the children were presented with a series of everyday social contexts and then asked to use language in a social way themselves. Though Ornaghi et al. (2011) also trained children in understanding and actively using mental state language, the present intervention was twice as short, yet positive effects were nonetheless obtained. The results of this study thus help to highlight the important role of language-based conversational training in improving preschooler's pragmatic abilities. Although this area has thus been relatively neglected in the research, we know that pragmatic competence is not only a key factor in children's language acquisition and socio-communicative development but also important for their general everyday functioning and wellbeing (Matthews et al., 2018; Norbury, 2014; Stephens & Matthews, 2014). In the context of our study, and in comparison with other developmental measures, the pragmatic test has been shown to be a highly sensitive tool that successfully captures children's learning outcomes with respect to social cognition measures. We thus suggest that researchers and practitioners include a comprehensive measure of children's pragmatic skills in studies that seek to assess mental state training outcomes in preschool children.

From a theoretical point of view, the results of the present study shed light on the complex relation between ToM and pragmatics, which is still open to debate (Bosco, Parola et al., 2018; Bosco, Tirassa, & Gabbatore, 2018; Sperber & Wilson, 2002b; Westra & Carruthers, 2017). Though it is indisputable that pragmatics and ToM are conceptually related (Cummings, 2015), the specific nature of this relationship is still unclear. As pointed out in Lecce, Ronchi, Del Sette, Bischetti, and Bambini (2019), divergent theoretical perspectives about the link between the two have emerged. One approach is a pragmatic account of ToM recently proposed by Frank

(2018) and Westra and Carruthers (2017), according to which ToM is regarded as pragmatic in nature. According to another approach (Sperber & Wilson, 2002a), the relationship between pragmatics and ToM stems from the fact that pragmatics requires the attribution of intentions to speakers and implies reasoning about their mental states. As a result, pragmatics is regarded as a sub-module of ToM. However, a view proposed by Bosco, Tirassa et al. (2018) argues that ToM and pragmatics may overlap in functioning but are not identical: providing compelling empirical evidence from developmental and clinical research, these authors claim that ToM alone cannot explain performance on pragmatic tasks.

In the current study, the two abilities were measured using two different tasks: ToM was assessed with false belief tasks and pragmatic abilities were assessed with a set of social scenarios. Our results showed that mental state training only led to improved pragmatic scores but had no influence on false belief performance. This finding highlights the distinct outcomes brought about by training on ToM and pragmatics, adding new evidence in support of the notion that ToM and pragmatics might be separable capacities, in line with Bosco, Tirassa et al. (2018). In addition, our results show that mental state understanding, a capacity which has been primarily associated with ToM, can foster the ability to interact in a socially appropriate way, an ability which was not directly trained. Thus, the observed development of the ability to interact socially points towards a generalization effect of the intervention. The training on mental states may have sharpened children's awareness of not only the mental states but also the emotions of others, enhancing their ability to interact socially. In other words, children's increased ability to pay attention to others' perspectives, which is necessary to achieve successful communication, combined with their increased ability to react in a socially appropriate way, resulted in improved performance on the pragmatic test. In sum, the results from the study provide indirect evidence that ToM and pragmatics constitute two distinct but complementary abilities and also show the generalization effect of mental state understanding on pragmatic abilities.

This study has a few limitations. First, it would have been of value to extend the duration of training in order to detect potential differences between the two experimental conditions. As noted above, the posttest scores of the two training groups differed as indicated by the larger effect sizes in the embodied group for the pragmatic competence and emotion understanding variables; however, the difference was not statistically significant. Based on the difference in effect size in the present study, one might expect more prolonged training to lead the embodied group to outperform the non-embodied group. Second, it would have been useful to have the children perform the same battery of tests again several months after the posttest to test the longevity of the intervention effects. Future research will be needed to confirm and expand on the present findings.

Nevertheless the present study highlights the important role of conversational classroom interventions in improving pragmatic abilities, whether accompanied by embodied communicative behaviors or not. These results are of special relevance to educational practices since the sort of conversational training described here can be easily applied to a variety of preschool classroom contexts, as well as laying the foundations for sociocognitive training-test batteries with atypically developed children. From an educational point of view, it is important to highlight that the current experimental conversational classroom intervention, both embodied and non-embodied, was carried out in the naturalistic context of schools and can thus be easily applied to a variety of educational settings. To this end, the training materials used in this study and accompanying instructions are freely available online at https://entrenemlesemocions.wordpress.com to aid preschool teachers to incorporate aspects of conversational training through embodiment into their classrooms. Moreover, the current findings reveal that from a methodological perspective it can be beneficial to adopt a multidimensional perspective and to include comprehensive measures of children's pragmatic competence to obtain a more complete picture of their developing socio-communicative abilities.

Author contributions

MP drafted the manuscript, collected and analyzed the data. IH, PP, and JH provided the initial design of the study, all authors contributed to the final design. IH, MP, and PP designed the experimental stimuli. MP and IH performed the piloting. MP, IH, and PP carried out the recruitment of the participants. All the authors contributed to the study's conceptualization, critically reviewed the manuscript for intellectual content, and contributed to manuscript edits.

Declaration of Competing Interest

None.

Acknowledgements

We are grateful to the preschool teachers and board of directors at the two participating schools in Barcelona, *Escola Antoni Brusi* and *Escola Bogatell*, and also to *Escola Verd* in Girona, for their participation in this study, as well as to the children who participated in this study and their parents. We would like to express our gratitude to research assistants Eva Castillo, Anna Massanas, and Ainhoa David for help with data collection. Very special thanks go to Anna Massanas for carrying out the training sessions. We are also very thankful to the committee members for the first author's PhD research plan defense, Wolfram Hinzen, Alfonso Igualada, and Núria Esteve-Gibert, for their valuable and insightful comments and suggestions. Furthermore, we would like to thank Joan Borràs-Comes for assistance with the statistical analyses and Alona Lukina and Valentin Sotskov for providing us with the python script we used to assign participants to groups. Thanks to the authors of training materials for permission to use the images. This research was funded by a research grant awarded by the Spanish Ministry of Science, Innovation, and Universities (PGC2018-097007-B-I00, 'Multimodal Language Learning (MLL): Prosodic and Gestural Integration in Pragmatic and Phonological Development'), and by a grant awarded

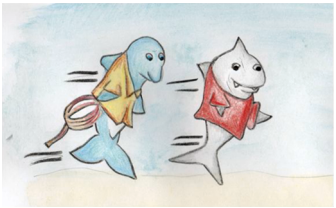
by the Generalitat de Catalunya (2017 SGR 971) to the Prosodic Studies Group. The second author was supported during the preparation of this manuscript by a postdoctoral research fellowship from the University of Zurich's URPP Language and Space. The first author also acknowledges a Novice Researcher grant from the Generalitat de Catalunya (ref. 2020 FI_B2 00192).

Appendix A. Sample story from the training materials in English translation

Story 8: ***Sara sea-turtle*** (target terms: *believing, getting upset*).

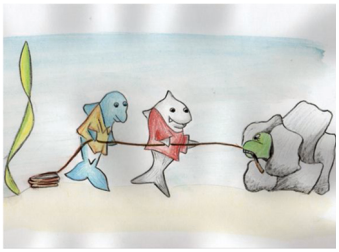
One day, Theo Shark was swimming along the bottom of the sea. Suddenly, he heard a voice calling for help and he thought: "Someone is in trouble or in danger". He looked around and decided to try to find that person and help them.

He tried to figure out where the voice was coming from by swimming about in all directions. When he came near some rocks he could hear the voice more strongly. He went right up to the rocks and saw an opening between two of them. He looked in and saw Sara Sea-Turtle who was crying. Theo asked her: "What happened? Why are you crying?".
Sara, upset and frightened, said: "I was gathering food for my little babies and I saw some delicious shellfish in this hole. I thought that my little ones would love them. I noticed that the hole was a bit small, but I decided to get in anyhow and now I'm stuck inside here and can't get out. Please help me!".
Theo replied: "Of course, don't worry, stay calm. I'll go and call my friend Jack Dolphin; he's sure to know how to get you out of there".



So Theo swum as fast as he could towards Jack's house. When he got there, he explained the situation to his friend, and Jack immediately had an idea: he decided that they would take a very strong rope with them to pull the sea-turtle out of the hole.

Together, Theo and Jack went back to Sara Sea-Turtle and when she saw them she felt a lot calmer because she know that two friends would do their best to get her out of the hole.
Theo Shark said: "We're back, my friend Jack has brought a rope; now you've got to hold it tightly between your teeth and we'll try to pull you out".
Sara, feeling much better now, said: "OK". She thought to herself: "I've got to try as hard as I can. I really hope it works!".



Theo and Jack began to pull as strongly as they could, but no matter how hard they tried, they could not get Sara out. "This isn't working", said Jack. "She's too stuck. We won't be able to get her out on our own."
So they decided to go and ask some other fish to help them.
"The more of us there are, the harder we will be able to pull!", thought Jack.

After a while, Charlie Balloon Fish, Amadeus Hammerhead, Sally Eel and even Diego Seahorse came to help too. When Sara saw so many other fish coming to help her she felt better straight away and tried even harder to free herself.

(*continued on next page*)

(*continued*)

They all pulled on the rope as hard as they could, and after a number of attempts, the sea-turtle finally popped out of the hole she had been stuck in. Delighted to be free at last, she thanked Theo, Jack and all the other fish, before rushing home to feed her babies.
From that day onwards, Sara never went through narrow openings again in case she got stuck once more.

References

Adams, C., Lockton, E., Freed, J., Gaile, J., Earl, G., McBean, K., … Law, J. (2012). The Social Communication Intervention Project: A randomized controlled trial of the effectiveness of speech and language therapy for school-age children who have pragmatic and social communication problems with or without autism spectrum disorder. *International Journal of Language & Communication Disorders, 47*(3), 233–244. https://doi.org/10.1111/j.1460-6984.2011.00146.x

Alonso-Alberca, N., Vergara, A. I., Fernández-Berrocal, P., Johnson, S. R., & Izard, C. E. (2012). The adaptation and validation of the emotion matching task for preschool children in Spain. *International Journal of Behavioral Development, 36*(6), 489–494. https://doi.org/10.1177/0165025412462154

Appleton, M., & Reddy, V. (1996). Teaching three year-olds to pass false belief tests: A conversational approach. *Social Development, 5*(3), 275–291. https://doi.org/10.1111/j.1467-9507.1996.tb00086.x

Armstrong, M., Esteve-Gibert, N., Hübscher, I., Igualada, A., & Prieto, P. (2018). Developmental and cognitive aspects of children's disbelief comprehension through intonation and facial gesture. *First Language, 38*(6), 596–616. https://doi.org/10.1177/0142723718789278

Astington, J. W., & Gopnik, A. (1988). Knowing you've changed your mind: Children's understanding of representational change. In J. W. Astington, P. L. Harris, & D. R. Olson (Eds.), *Developing theories of mind* (pp. 193–206). New York, NY, US: Cambridge University Press.

Astington, J. W., & Jenkins, J. M. (1999). A longitudinal study of the relation between language and theory-of-mind development. *Developmental Psychology, 35*(5), 1311–1320. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/10493656.

Astington, J. W., & Peskin, J. (2004). Meaning and use: Children's acquisition of the mental lexicon. *The Development of the Mediated Mind: Sociocultural Context and Cognitive Development*, 59–78.

Aureli, T., Spinelli, M., Fasolo, M., Garito, M. C., Perucchini, P., & D'Odorico, L. (2017). The pointing–vocal coupling progression in the first half of the second year of life. *Infancy, 22*(6), 801–818. https://doi.org/10.1111/infa.12181

Bambini, V., Arcara, G., Bechi, M., Buonocore, M., Cavallaro, R., & Bosia, M. (2016). The communicative impairment as a core feature of schizophrenia: Frequency of pragmatic deficit, cognitive substrates, and relation with quality of life. *Comprehensive Psychiatry, 71*, 106–120. https://doi.org/10.1016/j.comppsych.2016.08.012

Baron-Cohen, S., Leslie, A. M., & Frith, U. (1985). Does the autistic child have a "theory of mind"? *Cognition, 21*(1), 37–46.

Bartz, D. T. (2017). *Young children's meta-ignorance*. PhD Thesis. Harvard University.

Bates, D., Mächler, M., Bolker, B. M., & Walker, S. C. (2015). Fitting linear mixed-effects models using lme4. *Journal of Statistical Software, 67*(1), 1–48. https://doi.org/10.18637/jss.v067.i01

Bates, E., Camaioni, L., & Volterra, V. (1975). The acquisition of performatives prior to speech. *Merrill-Palmer Quarterly of Behavior and Development, 21*(3), 205–226.

Behne, T., Carpenter, M., & Tomasello, M. (2005). One-year-olds comprehend the communicative intentions behind gestures in a hiding game. *Developmental Science, 8*(6), 492–499. https://doi.org/10.1111/j.1467-7687.2005.00440.x

Berman, J. M. J., Chambers, C. G., & Graham, S. A. (2010). Preschoolers' appreciation of speaker vocal affect as a cue to referential intent. *Journal of Experimental Child Psychology, 107*(2), 87–99. https://doi.org/10.1016/j.jecp.2010.04.012

Berman, J. M. J., Chambers, C. G., & Graham, S. A. (2016). Preschoolers' real-time coordination of vocal and facial emotional information. *Journal of Experimental Child Psychology, 142*, 391–399. https://doi.org/10.1016/j.jecp.2015.09.014

Berman, J. M. J., Graham, S. A., Callaway, D., & Chambers, C. G. (2013). Preschoolers use emotion in speech to learn new words. *Child Development, 84*(5), 1791–1805. https://doi.org/10.1111/cdev.12074

Berman, J. M. J., Graham, S. A., & Chambers, C. G. (2013). Contextual influences on children's use of vocal affect cues during referential interpretation. *The Quarterly Journal of Experimental Psychology, 66*(4), 705–726. https://doi.org/10.1080/17470218.2012.713367

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

Bosco, F. M., Parola, A., Angeleri, R., Galetto, V., Zettin, M., & Gabbatore, I. (2018). Improvement of communication skills after traumatic brain injury: The efficacy of the cognitive pragmatic treatment program using the communicative activities of daily living. *Archives of Clinical Neuropsychology, 33*(7), 875–888. https://doi.org/10.1093/arclin/acy041

Bosco, F. M., Tirassa, M., & Gabbatore, I. (2018). Why pragmatics and Theory of Mind do not (completely) overlap. *Frontiers in Psychology, 9*, 1453. https://doi.org/10.3389/fpsyg.2018.01453

Camaioni, L., Perucchini, P., Bellagamba, F., & Colonnesi, C. (2004). The role of declarative pointing in developing a theory of mind. *Infancy, 5*(3), 291–308. https://doi.org/10.1207/s15327078in0503_3

Carpenter, M., Akhtar, N., & Tomasello, M. (1998). Fourteen-through 18-month-old infants differentially imitate intentional and accidental actions. *Infant Behavior and Development, 21*(2), 315–330.

Carrow-Woolfolk, E. (1999). *CASL: Comprehensive assessment of spoken language*. Circle Pines, MN: American Guidance Services.

Carrow-Woolfolk, E. (2017). *Comprehensive assessment of spoken language* (2nd ed. (casl-2)). Torrance, CA: Western Psychological Services.

Casillas, M., & Frank, M. C. (2012). Cues to turn boundary prediction in adults and preschoolers. In S. Brown-Schmidt, J. Ginzburg, & S. Larsson (Eds.), *Proceedings of SemDial 2012 (SeineDial): The 16th workshop on the semantics and pragmatics of dialogue* (pp. 61–69). Paris: Université Paris-Diderot.

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. https://doi.org/10.1111/1467-7687.00102

Colletta, J. M., Guidetti, M., Capirci, O., Cristilli, C., Demir, O. E., Kunene-Nicolas, R. N., & Levine, S. C. (2015). Effects of age and language on co-speech gesture production: An investigation of French, American, and Italian children's narratives. *Journal of Child Language, 42*(1), 122–145. https://doi.org/10.1017/S0305000913000585

Conte, E., Ornaghi, V., Grazzani, I., Pepe, A., & Cavioni, V. (2019). Emotion knowledge, theory of mind, and language in young children: Testing a comprehensive conceptual model. *Frontiers in Psychology, 10*, 2144. https://doi.org/10.3389/fpsyg.2019.02144

Cortés, A. I. R., Cobos, F. J. M., & Tarbox, J. (2018). Teaching children to create metaphorical expressions. *International Journal of Psychology and Psychological Therapy, 18*(1), 27–38. Retrieved from http://www.

Cummings, L. (2015). Theory of mind in utterance interpretation: The case from clinical pragmatics. *Frontiers in Psychology*, (August 26)https://doi.org/10.3389/fpsyg.2015.01286. Frontiers.

Davies, C., McGillion, M., Rowland, C., & Matthews, D. (2019). Can inferencing be trained in preschoolers using shared book-reading? A randomised controlled trial of parents' inference-eliciting questions on oral inferencing ability. *Journal of Child Language*, 1–25. https://doi.org/10.1017/S0305000919000801

de Villiers, J. (1995). Steps in the mastery of sentence complements. In *Paper Presented at the Biennial Meeting of the Society for Research in Child Development*.

de Villiers, J. (2000). Language and theory of mind: What are the developmental relationships? In S. Baron-Cohen, H. Tager-Flusberg, & D. J. Cohen (Eds.), *Understanding other minds: Perspectives from developmental cognitive neuroscience* (pp. 83–123). New York, NY, US: Oxford University Press.

de Villiers, J., & De Villiers, P. A. (2000). Linguistic determinism and the understanding of false beliefs. In P. Mitchell, & K. J. Riggs (Eds.), *Children's reasoning and the mind* (pp. 191–228). Hove, UK: Psychology Press.

de Villiers, J., & Pyers, J. E. (2002). Complements to cognition: A longitudinal study of the relationship between complex syntax and false-belief-understanding. *Cognitive Development, 17*(1), 1037–1060. https://doi.org/10.1016/S0885-2014(02)00073-4

Domaneschi, F., & Bambini, V. (2020). Pragmatic competence. In E. Fridland, & C. Pavese (Eds.), *Routledge handbook of skill and expertise*. Springer. https://doi.org/10.31234/osf.io/q8msa.

Esteban, M., Sidera, F., Serrano, J., Amadó, A., & Rostan, C. (2010). Improving social understanding of preschool children: Evaluation of a training program. *Electronic Journal of Research in Education Psychology, 8*(2), 841–860. https://doi.org/10.25115/ejrep.v8i21.1386

Esteve-Gibert, N., & Prieto, P. (2018). Early development of the prosody-meaning interface. In P. Prieto, & N. Esteve-Gibert (Eds.), *The development of prosody in first language acquisition* (pp. 227–246). Amsterdam: John Benjamins.

Esteve-Gibert, N., Lœvenbruck, H., Dohen, M., & D'Imperio, M. P. (2017). The use of prosody and gestures for the production of contrastive focus in French-speaking 4 and 5 year olds. In *Proceedings of the Workshop on Abstraction, Diversity and Speech Dynamics*.

Esteve-Gibert, N., Prieto, P., & Liszkowski, U. (2017). Twelve-month-olds understand social intentions based on prosody and gesture shape. *Infancy, 22*(1), 108–129. https://doi.org/10.1111/infa.12146

Farroni, T., Csibra, G., Simion, F., & Johnson, M. H. (2002). Eye contact detection in humans from birth. *Proceedings of the National Academy of Sciences of the United States of America, 99*(14), 9602–9605. https://doi.org/10.1073/pnas.152159999

Frank, C. K. (2018). Reviving pragmatic theory of theory of mind. *AIMS Neuroscience, 5*(2), 116–131. https://doi.org/10.3934/Neuroscience.2018.2.116

Frota, S., & Butler, J. (2018). Early development of intonation. In P. Prieto, & N. Esteve-Gibert (Eds.), *The development of prosody in first language acquisition* (pp. 145–164). Amsterdam: John Benjamins. https://doi.org/10.1075/tilar.23.08fro.

Gabbatore, I., Sacco, K., Angeleri, R., Zettin, M., Bara, B. G., & Bosco, F. M. (2015). Cognitive pragmatic treatment: A rehabilitative program for traumatic brain injury individuals. *The Journal of Head Trauma Rehabilitation, 30*(5), E14–E28. https://doi.org/10.1097/HTR.0000000000000087

Ganzeboom, H. B. G., De Graaf, P. M., Treiman, D. J., & De Leeuw, J. (1992). A standard international socio-economic index of occupational status. *Social Science Research, 21*, 1–56.

Goldin-Meadow, S. (1998). The development of gesture and speech as an integrated system. *New Directions for Child and Adolescent Development, 1998*(79), 29–42. https://doi.org/10.1093/acprof

Goldin-Meadow, S. (2014). Widening the lens: What the manual modality reveals about language, learning and cognition. *Philosophical Transactions of the Royal Society B: Biological Sciences, 369*(1651), 20130295. https://doi.org/10.1098/rstb.2013.0295

Goldin-Meadow, S., & Wagner, S. M. (2005). How our hands help us learn. *Trends in Cognitive Sciences, 9*(5), 230–241. https://doi.org/10.1016/j.tics.2005.03.006

González-Fuente, S. (2017). *Audiovisual prosody and verbal irony*. PhD Thesis. Barcelona: Universitat Pompeu Fabra.

Gopnik, A., & Astington, J. W. (1988). Children's understanding of representational change and its relation to the understanding of false belief and the appearance-reality distinction. *Child Development, 59*(1), 26–37. https://doi.org/10.1111/j.1467-8624.1988.tb03192.x

Graziano, M. (2014). The development of two pragmatic gestures of the so-called Open hand supine family in Italian children. In M. Seyfeddinipur, & M. Gullberg (Eds.), *From gesture in conversation to visible action as utterance: Essays in honor of Adam Kendon* (pp. 311–330). Amsterdam: John Benjamins.

Grazzani, I. G., & Ornaghi, V. (2011). Emotional state talk and emotion understanding: A training study with preschool children. *Journal of Child Language, 38*(5), 1124–1139. https://doi.org/10.1017/S0305000910000772

Grünloh, T., & Liszkowski, U. (2015). Prelinguistic vocalizations distinguish pointing acts. *Journal of Child Language, 42*(06), 1312–1336. https://doi.org/10.1017/S0305000914000816

Guajardo, N. R., & Watson, A. C. (2002). Narrative discourse and Theory of Mind development. *The Journal of Genetic Psychology, 163*(3), 305–325. https://doi.org/10.1080/00221320209598686

Hale, C. M., & Tager-Flusberg, H. (2003). The influence of language on theory of mind: A training study. *Developmental Science, 6*(3), 346–359. https://doi.org/10.1111/1467-7687.00289

Hale, C. M., & Tager-Flusberg, H. (2005). Social communication in chidren with autism: The relationship between theory of mind and discourse development. *Autism, 9*(2), 157–178. https://doi.org/10.1177/1362361305051395

Happé, F. (1995). The role of age and verbal ability in the theory of mind task performance of subjects with autism. *Child Development, 66*(3), 843–855. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/7789204.

Harris, P. L. (2006). Social cognition. In W. Damon (Ed.), *Handbook of child psychology* (pp. 811–857). New York: Wiley. https://doi.org/10.1002/9780470147658.chpsy0219.

Harvey, P. D., & Penn, D. (2010). February). Social cognition: The key factor predicting social outcome in people with Schizophrenia? *Psychiatry*. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/20376275.

Henry, J. D., Phillips, L. H., Ruffman, T., & Bailey, P. E. (2013). A meta-analytic review of age differences in theory of mind. *Psychology and Aging, 28*(3), 826–839. https://doi.org/10.1037/a0030677

Hoicka, E., & Wang, S. (2011). Fifteen-month-old infants match vocal cues to intentional actions. *Journal of Cognition and Development, 12*(3), 299–314. https://doi.org/10.1080/15248372.2010.542215

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

Hübscher, I., Esteve-Gibert, N., Igualada, A., & Prieto, P. (2017). Intonation and gesture as bootstrapping devices in speaker uncertainty. *First Language, 37*(1), 24–41. https://doi.org/10.1177/0142723716673953

Hübscher, I., Prieto, P., & Wagner, L. (2020). Three-year-olds infer polite stance from intonation and facial cues. *Journal of Politeness Research Language Behaviour and Culture, 16*(1), 85–110.

Hübscher, I., Garufi, L., & Prieto, P. (2019). The development of polite stance in preschoolers: how prosody, gesture and body cues pave the way. *Journal of Child Language*, (46), 825–46862. https://doi.org/10.1017/S0305000919000126

Hübscher, I., Vincze, L., & Prieto, P. (2019). Children's signaling of their uncertain knowledge state: Prosody, face and body cues come first. *Journal of Language Learning and Development, 15*(4), 366–389. https://doi.org/10.1080/15475441.2019.1645669
Hughes, C., & Leekam, S. (2004). What are the links between theory of mind and social relations? Review, reflections and new directions for studies of typical and atypical development. *Social Development, 13*(4), 590–619. https://doi.org/10.1111/j.1467-9507.2004.00285.x
Izard, C. E., Haskins, F. W., Schultz, D., Trentacosta, C. J., & King, K. A. (2003). *Emotion matching task*. Newark, DE: University of Delaware.
Kahana-Kalman, R., & Walker-Andrews, A. S. (2001). The role of person familiarity in young infants' perception of emotional expressions. *Child Development, 72*(2), 352–369. https://doi.org/10.1111/1467-8624.00283
Khu, M., Chambers, C. G., & Graham, S. A. (2018). When you're happy and I know it: Four-year-olds' emotional perspective taking during online language comprehension. *Child Development, 89*(6), 2264–2281. https://doi.org/10.1111/cdev.12855
Kim, S., Paulus, M., Sodian, B., & Proust, J. (2016). Young children's sensitivity to their own ignorance in informing others. *PloS One, 11*(3), e0152595. https://doi.org/10.1371/journal.pone.0152595
Kunda, Z. (1999). Social cognition: Making sense of people. *Contemporary Sociology, 29*(6), 836. https://doi.org/10.2307/2654104
Laval, V., & Bert-Erboul, A. (2005). French-speaking children's understanding of sarcasm: The role of intonation and context. *Journal of Speech Language and Hearing Research, 48*(3), 610–620. https://doi.org/10.1044/1092-4388(2005/042)
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. https://doi.org/10.1016/j.jecp.2011.04.011
Lecce, S., Ronchi, L., Del Sette, P., Bischetti, L., & Bambini, V. (2019). Interpreting physical and mental metaphors: Is theory of mind associated with pragmatics in middle childhood? *Journal of Child Language, 46*(2), 393–407. https://doi.org/10.1017/S030500091800048X
Lecce, S., Bianco, F., Demicheli, P., & Cavallini, E. (2014). Training preschoolers on first-order false belief understanding: Transfer on advanced ToM skills and metamemory. *Child Development, 85*(6), 2404–2418. https://doi.org/10.1111/cdev.12267
Lecce, S., Caputi, M., & Pagnin, A. (2014). Long-term effect of theory of mind on school achievement: The role of sensitivity to criticism. *European Journal of Developmental Psychology, 11*(3), 305–318. https://doi.org/10.1080/17405629.2013.821944
Lenth, R., Singmann, H., Love, J., Buerkner, P., & Herve, M. (2018). Emmeans: Estimated marginal means, aka least-squares means. 2018. *R Package Version, 1*(4).
Llanes-Coromina, J., Vilà-Giménez, I., Kushch, O., Borràs-Comes, J., & Prieto, P. (2018). Beat gestures help preschoolers recall and comprehend discourse information. *Journal of Experimental Child Psychology, 172*, 168–188. https://doi.org/10.1016/j.jecp.2018.02.004
Lohmann, H., & Tomasello, M. (2003). The role of language in the development of false belief understanding: A training study. *Child Development, 74*(4), 1130–1144. https://doi.org/10.1111/1467-8624.00597
Lüdecke, D. (2016). *Sjmisc: Data transformation and labelled data utility functions*. Retrieved from https://mran.microsoft.com/snapshot/2017-04-22/web/packages/sjmisc/sjmisc.pdf.
Matthews, D. (2014). In D. Matthews (Ed.), *Pragmatic development in first language acquisition*. Amsterdam: John Benjamins Publishing Company. https://doi.org/10.2307/23093720.
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. https://doi.org/10.1080/15475441.2018.1455584
McNeil, N. M., Alibali, M. W., & Evans, J. L. (2000). The role of gesture in children's comprehension of spoken language: Now they need it, now they don't. *Journal of Nonverbal Behavior, 24*(2), 131–150. https://doi.org/10.1023/A:1006657929803
Melot, A. M., & Angeard, N. (2003). Theory of mind: Is training contagious? *Developmental Science, 6*(2), 178–184. https://doi.org/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. https://doi.org/10.1111/j.1467-8624.2007.01018.x
Norbury, C. (2014). Practitioner review: Social (pragmatic) communication disorder conceptualization, evidence and clinical implications. *Journal of Child Psychology and Psychiatry, and Allied Disciplines, 55*(3), 204–216. https://doi.org/10.1111/jcpp.12154
Oller, D. K., Buder, E. H., Ramsdell, H. L., Warlaumont, A. S., Chorna, L., & Bakeman, R. (2013). Functional flexibility of infant vocalization and the emergence of language. *Proceedings of the National Academy of Sciences, 110*(16), 6318–6323.
Ornaghi, V., & Grazzani, I. G. (2020). *Lo sviluppo della comprensione sociale nei bambini: un laboratorio linguistico per promuovere la teoria della mente e la comprensione delle emozioni*. Parma: Junior-Spaggiari.
Ornaghi, V., Brockmeier, J., & Grazzani, 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. https://doi.org/10.1080/15248372.2011.563487
Ornaghi, V., Pepe, A., & Grazzani, I. G. (2016). False-belief understanding and language ability mediate the relationship between emotion comprehension and prosocial orientation in preschoolers. *Frontiers in Psychology, 7*, 1534. https://doi.org/10.3389/fpsyg.2016.01534
Parsons, L., Cordier, R., Munro, N., Joosten, A., & Speyer, R. (2017). A systematic review of pragmatic language interventions for children with autism spectrum disorder. *PloS One, 12*(4), e0172242. https://doi.org/10.1371/journal.pone.0172242
Pelletier, J., & Astington, J. W. (1998). *Metacognitive vocabulary test*. Institute of Child Study, University of Toronto. Unpublished.
Perner, J. (1991). *Understanding the representational mind*. Cambridge, MA: The MIT Press.
Perner, J., & Roessler, J. (2012). From infants' to children's appreciation of belief. *Trends in Cognitive Sciences, 16*(10), 519–525. https://doi.org/10.1016/j.tics.2012.08.004
Perner, J., Leekam, S. R., & Wimmer, H. (1987). Three-year-olds' difficulty with false belief: The case for a conceptual deficit. *The British Journal of Developmental Psychology, 5*(2), 125–137. https://doi.org/10.1111/j.2044-835X.1987.tb01048.x
Peterson, C. C., O'Reilly, K., & Wellman, H. M. (2016). Deaf and hearing children's development of theory of mind, peer popularity, and leadership during middle childhood. *Journal of Experimental Child Psychology, 149*, 146–158. https://doi.org/10.1016/j.jecp.2015.11.008
Pexman, P. M. (2019). The role of embodiment in conceptual development. *Language, Cognition and Neuroscience, 34*(10), 1274–1283. https://doi.org/10.1080/23273798.2017.1303522
Pexman, P. M., Reggin, L., & Lee, K. (2019). Addressing the challenge of verbal irony: Getting serious about sarcasm training. *Languages, 4*(2), 23. https://doi.org/10.3390/languages4020023
Phelps-Terasaki, D., & Phelps-Gunn, T. (2007). *The test of pragmatic language* (2nd ed.). Austin, TX: Pro-Ed.
Pons, F., Harris, P. L., & de Rosnay, M. (2004). Emotion comprehension between 3 and 11 years: Developmental periods and hierarchical organization. *The European Journal of Developmental Psychology, 1*(2), 127–152. https://doi.org/10.1080/17405620344000022
Premack, D., & Woodruff, G. (1978). Does the chimpanzee have a theory of mind? *The Behavioral and Brain Sciences, 1*(4), 515–526. https://doi.org/10.1017/S0140525X00076512
Pronina, M., Hübscher, I., Vilà-Giménez, I., & Prieto, P. (2019). A new tool to assess pragmatic prosody in children: Evidence from 3- to 4-year-olds. In *Proceedings of the 19th International Congress of Phonetic Sciences*.
R Core Team. (2019). *R: A language and environment for statistical computing*. Retrieved from. Vienna, Austria: Foundation for Statistical Computing https://www.r-project.org/.
Rusiewicz, H. L., & Esteve-Gibert, N. (2018). Set in time: Prosody-Gesture coordination. In P. Prieto, & N. Esteve-Gibert (Eds.), *The development of prosody in first language acquisition* (pp. 103–124). Amsterdam: John Benjamins.
Saborit Mallol, C., Julián Marzá, J. P., & Navarro Lizandra, J. L. (2005). *ELI : l'avaluació del llenguatge infantil*. Castelló de la Plana: Publicacions de la Universitat Jaume I.
Sakkalou, E., & Gattis, M. (2012). Infants infer intentions from prosody. *Cognitive Development, 27*(1), 1–16. https://doi.org/10.1016/j.cogdev.2011.08.003
Scheiner, E., Hammerschmidt, K., Jürgens, U., & Zwirner, P. (2002). Acoustic analyses of developmental changes and emotional expression in the preverbal vocalizations of infants. *Journal of Voice, 16*(4), 509–529. https://doi.org/10.1016/S0892-1997(02)00127-3

Sekine, K., & Kita, S. (2015). The parallel development of the form and meaning of two-handed gestures and linguistic information packaging within a clause in narrative. *Open Linguistics, 1*, 490–502.

Slade, L., & Ruffman, T. (2005). How language does (and does not) relate to theory of mind: A longitudinal study of syntax, semantics, working memory and false belief. *The British Journal of Developmental Psychology, 23*(1), 117–141. https://doi.org/10.1348/026151004X21332

Slaughter, V. (1998). Children's understanding of pictorial and mental representations. *Child Development, 69*(2), 321–332. https://doi.org/10.1111/j.1467-8624.1998.tb06191.x

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. https://doi.org/10.1111/j.1467-8624.1996.tb01898.x

Sperber, D., & Wilson, D. (2002a). Pragmatics, modularity and mind-reading. *Mind and Language, 17*(1), 3–33. https://doi.org/10.1111/1468-0017.00186

Sperber, D., & Wilson, D. (2002b). Relevance theory. *Handbook of pragmatics*. Blackwell.

Stephens, G., & Matthews, D. (2014). The communicative infant from 0–18 months. In *Pragmatic development in first language acquisition* (Vol. 10, pp. 13–36). Philadelphia: John Benjamins Publishing Company Amsterdam.

Szücs, M., & Babarczy, A. (2017). The role of theory of mind, grammatical competence and metapragmatic awareness in irony comprehension. In S. Assimakopoulos (Ed.), *Pragmatics at its interfaces* (Vol. 17, pp. 129–147). Berlin: Mouton de Gruyter. https://doi.org/10.1515/9781501505089-008.

Tager-Flusberg, H., & Joseph, R. M. (2012). How language facilitates the acquisition of false-belief understanding in children with autism. In J. W. Astington, & J. A. Baird (Eds.), *Why language matters for theory of mind* (pp. 298–318). New York: Oxford University Press. https://doi.org/10.1093/acprof:oso/9780195159912.003.0014.

Tenenbaum, H. R., Alfieri, L., Brooks, P. J., & Dunne, G. (2008). The effects of explanatory conversations on children's emotion understanding. *The British Journal of Developmental Psychology, 26*(2), 249–263. https://doi.org/10.1348/026151007X231057

Tomasello, M., Carpenter, M., & Liszkowski, U. (2007). A new look at infant pointing. *Child Development, 78*(3), 705–722. https://doi.org/10.1111/j.1467-8624.2007.01025.x

Wellman, H. M. (2014). *Making minds: How theory of mind develops*. Oxford University Press.

Wellman, H. M. (2018). Theory of mind: The state of the art. *The European Journal of Developmental Psychology, 15*(6), 728–755. https://doi.org/10.1080/17405629.2018.1435413

Wellman, H. M., & Liu, D. (2004). Scaling of theory-of-mind tasks. March 1. *Child development*. John Wiley & Sons, Ltd (10.1111). https://doi.org/10.1111/j.1467-8624.2004.00691.x.

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.

Westra, E., & Carruthers, P. (2017). *Pragmatic development explains the theory-of-Mind scale*. https://doi.org/10.1016/j.cognition.2016.10.021

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. https://doi.org/10.1016/0010-0277(83)90004-5