

AperTO - Archivio Istituzionale Open Access dell'Università di Torino

A broad assessment of theory of mind in adolescence: The complexity of mindreading

This is the author's manuscript

Original Citation:

Availability:

This version is available <http://hdl.handle.net/2318/141515> since

Published version:

DOI:10.1016/j.concog.2014.01.003

Terms of use:

Open Access

Anyone can freely access the full text of works made available as "Open Access". Works made available under a Creative Commons license can be used according to the terms and conditions of said license. Use of all other works requires consent of the right holder (author or publisher) if not exempted from copyright protection by the applicable law.

(Article begins on next page)

This Accepted Author Manuscript (AAM) is copyrighted and published by Elsevier. It is posted here by agreement between Elsevier and the University of Turin. Changes resulting from the publishing process - such as editing, corrections, structural formatting, and other quality control mechanisms - may not be reflected in this version of the text. The definitive version of the text was subsequently published in CONSCIOUSNESS AND COGNITION, 24 (1), 2014, 10.1016/j.concog.2014.01.003.

You may download, copy and otherwise use the AAM for non-commercial purposes provided that your license is limited by the following restrictions:

(1) You may use this AAM for non-commercial purposes only under the terms of the CC-BY-NC-ND license.

(2) The integrity of the work and identification of the author, copyright owner, and publisher must be preserved in any copy.

(3) You must attribute this AAM in the following format: Creative Commons BY-NC-ND license (<http://creativecommons.org/licenses/by-nc-nd/4.0/deed.en>), 10.1016/j.concog.2014.01.003

The publisher's version is available at:

<http://linkinghub.elsevier.com/retrieve/pii/S105381001400004X>

When citing, please refer to the published version.

Link to this full text:

<http://hdl.handle.net/2318/141515>

This full text was downloaded from iris - AperTO: <https://iris.unito.it/>

iris - AperTO

University of Turin's Institutional Research Information System and Open Access Institutional Repository

A broad assessment of theory of mind in adolescence: The complexity of mindreading

Francesca M. Bosco, Ilaria Gabbatore, Maurizio Tirassa

The aim of this research was to provide an articulated assessment of several different ToM components, namely first- vs. third-person, egocentric vs. allocentric, and first- vs. second-order ToM, in preadolescence and adolescence. Our expectations for the sample of 80 juveniles that participated in the research were that: (1) ToM abilities would improve with age; (2) participants would perform better at first-person than at third-person tasks; (3) participants would perform better at first-order than at second-order tasks; (4) girls will perform systematically better than boys. We also explored possible differences in performance (5) in the allocentric vs. the egocentric perspectives as well as (6) in the comprehension of different types of mental states, namely desires, beliefs and positive and negative emotions. Overall our expectations were confirmed. Our data confirmed that all ToM aspects we investigated keep maturing during preadolescence and adolescence.

1. Introduction

The phrase Theory of Mind (henceforth, ToM) was first introduced by Premack and Woodruff (1978) to refer to the capacity of ascribing mental states to oneself and the others and using this knowledge to predict and explain the relevant actions and behaviors. For various theoretical reasons, other labels were created in the following years to refer to this faculty, like mindreading (Baron-Cohen, 1995) or social cognition (Adolphs, 1999).

Initially ToM was implicitly considered a unitary, all-or-nothing faculty. With time, however, it has become clear that it has a more complex, articulated nature, opening the way to the possibility of decomposing it into different aspects or components. In particular, neuroscientific studies have identified a network of brain regions subserving social cognition, which is now commonly called the “social brain” (Brothers, 1990; Frith, 2007; Frith & Frith, 2010).

The goal of this research is to contribute to understanding how the different aspects of theory of mind develop through preadolescence and adolescence.

As a first thing, it is necessary to distinguish between *first-person* and *third-person ToM* (Nichols & Stich, 2003). To understand oneself and to understand another person appear to be different activities, mediated by different processes and recruiting different kinds of knowledge. This distinction is also supported by evidence that different brain circuitry is recruited when participants are asked to take a first- or a third-person perspective (Abu-Akel, 2003; Vogeley & Fink, 2003; Vogeley et al., 2001).

As regards third-person ToM, a difference is commonly drawn between *first-order* and *second-order ToM*. The former is the ability to grasp someone’s mental states (“John thinks that. . .”), while the latter requires to deal with nested representations,

that is to infer what someone thinks about a third person ("John thinks that Mary thinks that..."). Expectably, there is evidence that first-order tasks are easier than second-order ones to normally developing children (Wellman & Liu, 2004). On the average, children begin to solve the former at three or four years of age, at least according to the most pessimistic studies (Wimmer & Perner, 1983), and the latter at about seven (Perner & Wimmer, 1985).

A further distinction, proposed by Frith and de Vignemont (2005), is that between an *egocentric* and an *allocentric* perspective. In the former the mental states of other persons are represented in relation to the self, while in the latter they are represented independently from the self.

ToM-like reasoning appears to emerge in human beings during the second year of life (Bosco, Friedman, & Lesile, 2006; Onishi & Baillargeon, 2005), possibly founding on more primitive capabilities for shared intersubjectivity (Tirassa, Bosco, & Colle, 2006a, 2006b), and continues to develop at least through adolescence (Bosacki, 2000, 2003; Choudhury, Blakemore, & Charman, 2006; Goldstein & Winner, 2012; for a review see Sebastian, Viding, Williams, & Blakemore, 2010). A thorough understanding of the functioning of ToM during preadolescence and adolescence is particularly interesting since these phases of life are characterized by marked behavioral, hormonal and physical changes (Coleman & Hendry, 1999) as well as by the further maturation of aspects of cognitive functioning, like the so-called executive functions, that appear to be related to ToM (Dumontheil, Apperly, & Blakemore, 2010; Taylor, Barker, Heavey, & McHale, 2012; for a review see Apperly, Samson, & Humphreys, 2009).

During adolescence readiness toward the social environment outside the family gains a wholly new degree of independence, showing individual, emotional, social and cultural dynamics which are profoundly different to those of infancy. It is reasonable to think that ToM abilities and strategies would change correspondingly, adapting to the new social needs that the individual faces.

First- and third-person ToM reasoning appears to improve with age (Hatcher, Hatcher, Berlin, Okla, & Richards, 1990). Dumontheil et al. (2010) showed that the ability to adopt another agent's point of view grows from infancy through adolescence and further improves in adulthood. In general terms, it is not clear precisely when ToM reaches a final degree of maturation: Maylor, Moulson, Muncer, and Taylor (2002) studied the performance of young participants aged between 16 and 29 years old (mean age: 19 years old) at advanced first-person ToM tasks, including a subset of Strange stories (Happé, 1994; Happé, Brownell, & Winner, 1999) as well as new stories devised along similar lines: the participants to their study obtained a mean score of 4 against a maximum available of 7, thus showing no sign of a ceiling effect.

Neuroscience offers further evidence that the social brain is still developing during adolescence (Blakemore, den Ouden, Choudhury, & Frith, 2007; Shaw, Grosbras, Leonard, Pike, & Paus, 2012; for a review see Blakemore, 2008; Burnett, Sebastian, Coehen Kadosh, & Blakemore, 2011). Data collected with functional magnetic resonance imaging (fMRI) show that in this phase of life the social brain undergoes an increase in connectivity, synaptic reorganization, and a general structural development (Sowell et al., 2003). This appears to be related with various cognitive and emotional aspects of mental functioning, as well as with ToM abilities (Moriguchi, Ohnishi, Mori, Matsuda, & Komaki, 2007).

The aim of the research we present here was to conduct a behavioral assessment as broad as possible of theory of mind abilities in adolescence. We used a recently developed tool, the Theory of Mind Assessment Scale (Th.o.m.a.s.: Bosco, Colle, De Fazio, et al., 2009; see also Bosco, Capozzi, Colle, Marostica, & Tirassa, 2014; Castellino, Bosco, Marshall, Marshall, & Veglia, 2011; Chiavarino et al., 2014; Laghi et al., 2014). This is a semi-structured interview which provides a detailed profile of different facets of ToM abilities, namely first- vs. second-order, first- vs. third-person, egocentric vs. allocentric. It also explores different types of mental states involved in ToM (beliefs, desires, positive emotions, and negative emotions) and of causal relations between them, both in terms of the possess of knowledge about them and of the ability to govern their dynamics, that is to put such knowledge at use.

We also administered the Strange Stories, which is another advanced ToM task (Happé, 1994; Italian version by Mazzola & Camaioni, 2002).

Based upon the available literature, thus, our expectations for the sample of preadolescent and adolescents that participated in the research were that:

1. There will be an improvement of ToM abilities with age.
2. Participants will perform better at first-person than at third-person tasks.
3. Participants will perform better at first-order than at second-order tasks.
4. Girls will perform better than boys in all ToM tasks: this expectation is rooted in one of the few certainties that can be gathered from the scarce empirical literature on ToM in adolescence (Bosacki & Astington, 1999).

For explorative purpose we also wanted to investigate:

5. Whether the participants will perform differently when taking an allocentric than an egocentric perspective.
6. Whether the participants will differently deal with different types of mental states, namely beliefs, desires, positive emotions, and negative emotions.

4. METHOD

2.1. Participants

The participants were eighty preadolescents and adolescents, 40 females and 40 males, ranging from 11 to 17 years of age (notated as “years;months”) ($M = 14;5$; $SD = 2;3$). Their education level ranged from 5 to 10 years ($M = 8;1$; $SD = 2;0$). They belonged to four age groups: 11;00–11;11 years ($M = 11;4$; $SD = 0;3$), 13–13;11 years ($M = 13;7$; $SD = 0;3$), 15–15;11 years ($M = 15;6$; $SD = 0;4$), and 17–17;11 years ($M = 17;4$; $SD = 0;4$). The groups had equal size and included an equal number of males and females.

All the participants were recruited from summer schools in Piedmont or via personal contacts. They were all native Italian speakers, with no previous history of significant neurological and/or psychiatric disorders or substance abuse.

All the subjects participated voluntarily in the study; they, as well as their parents, were informed about the research procedure and gave their informed consent. The Bio-ethical committee of the University of Turin approved the study.

2.2. Materials and procedures

2.2.1. The Theory of Mind Assessment Scale (Th.o.m.a.s.)

Th.o.m.a.s. (Bosco, Colle, De Fazio, et al., 2009) is a semi-structured interview aimed at assessing a subject's theory of mind. It consists of 39 open-ended questions that leave the interviewee free to express and articulate her thought. When not spontaneously provided by the interviewee, the interviewer may specifically ask for real-world examples to enrich and contextualize the answer.

The interview is originally in Italian. The questions of which it is composed (see Appendix A) are organized along four scales, each focusing on one of the knowledge domains in which a person's ToM may manifest itself:

- *Scale A, I-Me*. This scale investigates the interviewee's knowledge of her own mental states, i.e. first-person ToM in an egocentric perspective. The viewpoint of the questions is centered on the interviewee (I) reflecting on her own mental states (Me); e.g., *Do you ever experience emotions that make you feel good?*
- *Scale B, Other-Self*. This scale investigates the knowledge that, according to the interviewee, the other persons have of their own mental states, independently of the interviewee's perspective, i.e., third-person ToM in an allocentric perspective. The viewpoint of the questions is centered on the other persons (Other) reflecting on their own mental states (Self); e.g., *Do the others try to fulfill their wishes?*
- *Scale C, I-Other*. This scale investigates the interviewee's alleged perception of the mental states of other persons.¹ The viewpoint of the questions is centered on the interviewee (I) reflecting on the others' mental states (Other); e.g., *Do you notice it when the others feel good?* This scale is similar to scale B in that both investigate third-person ToM; however, while the perspective there is centered on the others, here it is centered on the interviewee. In other words, here the subject is asked to take an egocentric perspective.
- *Scale D, Other-Me*. This scale investigates the knowledge that, from the interviewee's point of view, the others have of her mental states. The viewpoint of the questions is centered on the other persons (Other) reflecting on the mental states of the interviewee (Me); e.g., *Do the others notice it when you feel good?* This scale substantially is a second-order ToM task, since the abstract form of the questions is: *What do you think that the others think that you think?*

Based on independent theorizing about the most important types of mental states that an agent's cognitive architecture has to comprise (Bosco, Colle, & Tirassa, 2009; Tirassa, 1999; Tirassa & Bosco, 2008), and leaving aside intentions, the questions focus on the interviewee's perception of *epistemic states* like knowledge, beliefs and so on, *volitional states* like desires, and *emotions*, both *positive* and *negative*.

Since different types of causal or logical relation may occur between these mental states types and between them and the external world, each scale is divided into three subscales, respectively exploring one of three such types of relation:

- *Awareness*. This subscale investigates the interviewee's ability to perceive and differentiate beliefs, desires and emotions in herself and in the others. Recognizing different types of mental states is a logically necessary precondition to understanding their further links and relations.
- *Relation*. This subscale investigates the interviewee's ability to recognize the causal relations that hold between different mental states and between them and the world on the one hand, or the resulting behaviors on the other hand, e.g. *When you feel bad, do you feel you understand why?* Being capable of connecting and integrating different mental states and of understanding their reciprocal relations and bidirectional connections with perception and action is necessary to draw up an explanatory theory of the mind and of the social world.

¹ The interviewee is left free to interpret the word “other” as she prefers. If the answer is too general, however, the interviewer would propose that a specific person be identified as the focus of reflection, like a friend, a familiar, or the spouse or partner.

- *Realization*. This subscale investigates the interviewee's alleged ability to adopt effective strategies aimed at achieving a desired state. For example: *Do you succeed in getting what you want? How?* Acting adaptively in the inner world and in social contexts requires not only having a theory of the causal relations of mental states with one another and with the world, but also knowing how to put this knowledge to use, so as to appropriately and successfully act upon the mental states and behaviors of one's own and of the others.

In a graphic representation of the structure of the interview (see [Appendix B](#)), the four scales and their subscales are the columns of a table whose rows represent the types of mental states investigated. Thus, each cell of the table represents a specific intersection of two of the dimensions that the interview considers. Each question in its turn refers to a specific cell of the table, that is it encourages the interviewee to express her understanding of the relevant aspect of the activities of the mind.

For example, question [2]: *When you feel good, does that make any difference to you?* explores the ability to identify the interviewee's own positive emotions (dimensions investigated: Relation and Positive emotions). Question [29]: *Do you think you understand the others' wishes?* encourages the interviewee to express her understanding of the desires of the others (dimensions investigated: Awareness and Desires). The same type of considerations apply to each question. This structure is replicated for all four scales.

The interviewee may be asked to provide one or more examples of what she is saying whenever she does not do so spontaneously.

2.2.2. Standard ToM Tasks

In addition to Th.o.m.a.s., and during the same session, all the participants were also administered other ToM tests:

- *Strange Stories* ([Happé, 1994](#)). We chose a selection of six stories from the Italian version of the test ([Mazzola & Camaioni, 2002](#)).
- *The Ice Cream Van Story* ([Baron-Cohen, 1989](#)) and the *Burglar Story* ([Happé & Frith, 1994](#)). These tests have been devised to investigate second-order ToM; specifically, the ability to understand a false belief about the belief of another character.

2.3. Procedure and scoring

With the authorization of the interviewees, all the Th.o.m.a.s. interviews were tape-recorded and then transcribed. The transcriptions were rated by two independent judges: these were two research assistants who had not participated to the administration phase and worked separately and blind with respect to the hypotheses of the research as well as to the age of the various interviewees whose answers they were scoring.

The task of each judge was to assign each answer a score from 0 to 4, according to given rating criteria (see [Bosco, Colle, De Fazio, et al., 2009](#)), and to insert it in the relevant cell of the correction grid (see [Appendix B](#)).

Inter-rater reliability was evaluated on 30 randomly selected participants (15% of the total sample). The level of agreement between the scores assigned by the two judges was calculated using the Intraclass Correlation Coefficient (ICC), which provides a generalized measure of inter-rater concordance adjusted for chance agreement between measurements. The ICC was .70, which indicates a good inter-rater agreement ([Altman, 1991](#)).

Once this first stage of work was done, the judges discussed all the responses to which they had given different scores, until they reached a complete agreement.

The same judges also scored the other ToM tests, following the relevant criteria available in the literature, namely assigning 0 to an incorrect answer and 1 to a correct one.

Finally, we needed to rule out the hypothesis that an age-dependent increase in the participants' pragmatic abilities might affect the results of the ToM assessment. To this aim, we scored 30% of the interviews on a second grid for the evaluation of the communicative-pragmatic performance (see [Appendix C](#)), created in accordance with the criteria for pragmatic evaluation found in the relevant literature ([Bosco, Angeleri, Zuffranieri, Bara, & Sacco, 2012](#); [Penn, 1985](#); [Prutting & Kittchner, 1987](#)). These scores were assigned by a third research assistant, again blind to the goals and hypotheses of the research and to the demographic characteristics of the participants.

3. Results

3.1. Performance at Th.o.m.a.s. and age difference

First, we analyzed the general performance of the adolescents at Th.o.m.a.s. The ANOVA analysis revealed an effect of the age group ($F_{(3,76)} = 13.41$; $p < .001$; $\eta^2 = .35$) on the general performance at the interview.

The adolescents' performance at the Th.o.m.a.s. scales was investigated with a repeated measure ANOVA with one between-subjects factor (type of group, with four levels: 11, 13, 15, 17 years) and a within-subjects factor (type of scale, with four levels: scale A I-Me; scale B Other-Self; scale C Me-Other; scale D Other-Me) (see [Fig. 1](#)).

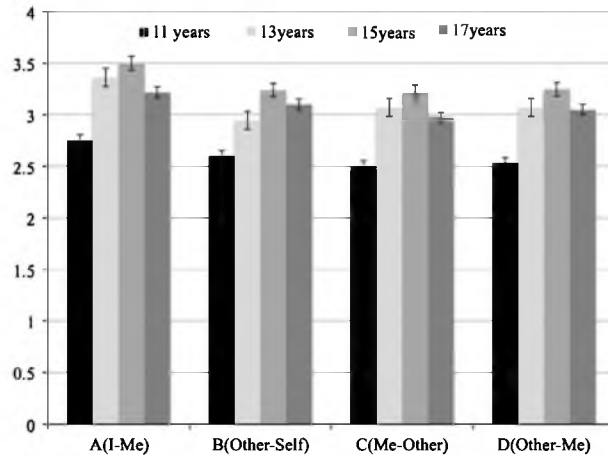


Fig. 1. Mean scores for each age group at Th.o.m.a.s. scales, with standard error bars.

The analysis revealed an effect of scale type ($F_{(3, 228)} = 21.95$; $p < .001$; $\eta^2 = .22$) and an effect of age group ($F_{(3,76)} = 13.41$; $p < .001$; $\eta^2 = .35$). We introduced a linear contrast which revealed a linear increase of scores depending on the age of the participants ($F_{(1,76)} = 45.96$; $p < .001$; $\eta^2 = .38$). A *post hoc* pairwise comparison revealed that the subjects scored higher to scale A (I-Me), which assesses first-order ToM, than to scales D (Other-Me, Bonferroni: $p < .001$), and C (Me-Other, Bonferroni: $p < .001$), both of which assess third-person ToM. No significant differences existed between scales C (Me-Other) and B (Other-Self), which focus respectively on the allocentric and the egocentric perspectives (Bonferroni: $p = 1$).

For exploratory purposes we evaluated the adolescents' performance at the Th.o.m.a.s. subscales with a repeated measure ANOVA with one between-subjects factor (type of group, with four levels: 11, 13, 15, 17 years) and a within-subjects factor (type of subscale, with three levels: Awareness, Relation, Realization) (see Fig. 2).

The analysis revealed an effect of subscale type ($F_{(2,152)} = 20.25$; $p < .001$; $\eta^2 = .21$) and an effect of subject group ($F_{(3,76)} = 13.31$; $p < .001$; $\eta^2 = .34$). A *post hoc* pairwise comparison revealed that the subjects scored higher at the Realization subscale than at both the Awareness (Bonferroni: $p = .01$) and the Relation (Bonferroni: $p < .001$) subscales. As regards the latter two, they scored better at Awareness than Relation (Bonferroni: $p = .003$). We also introduced a linear contrast which revealed a linear increase in scores depending on the age of the participants ($F_{(1,76)} = 9.20$; $p = .003$; $\eta^2 = .11$). However, as may be observed in Fig. 2, the differences between age groups did not increase by the same extent from each group age to the subsequent one; thus, to further investigate whether the adolescents' performance at Th.o.m.a.s. does increase with age, we performed a correlation between the scores obtained at Th.o.m.a.s. and the age of the participants, finding a significant one ($r = .466$; $p < .001$).

We investigated the adolescents' performance at the various types of mental states taken into account in Th.o.m.a.s. with a repeated measure ANOVA with one between-subjects factor (type of group, with four levels: 11, 13, 15, 17 years) and a within-subjects factor (type of subscale, with four levels: Beliefs, Desires, Positive emotions, Negative emotions) (see Fig. 3).

The analysis revealed an effect of dimension type ($F_{(3,228)} = 11.87$; $p < .001$; $\eta^2 = .13$) and an effect of subject group ($F_{(3,76)} = 14.32$; $p < .001$; $\eta^2 = .36$). We introduced a linear contrast which revealed a linear increase in scores depending

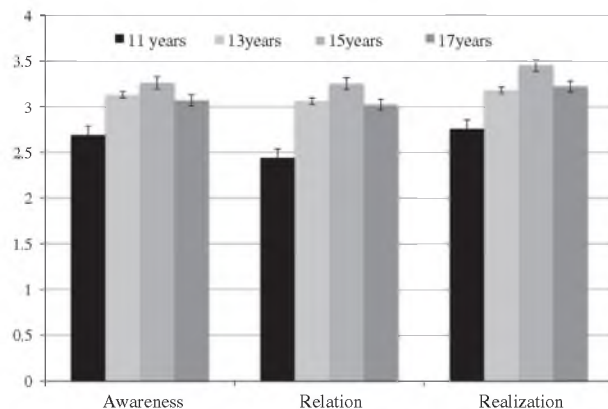
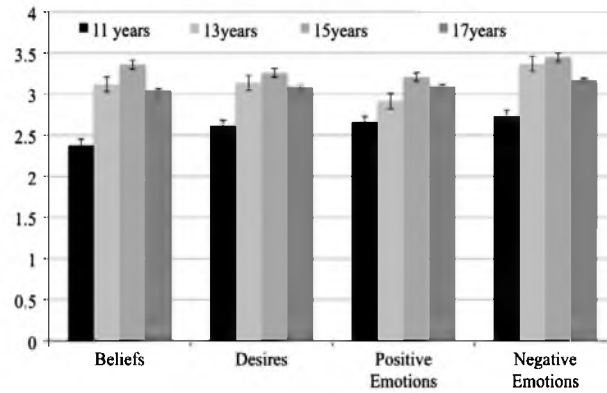


Fig. 2. Mean scores for each age group at Th.o.m.a.s. subscales, with standard error bars.



Mean scores for each age group at Th.o.m.a.s. dimensions, with standard error bars.

on the age of the participants ($F_{(1,76)} = 19.6$; $p < .001$; $\eta^2 = .20$). Furthermore, *post hoc* pairwise comparison revealed that the participants scored better at Negative emotions than at Beliefs (Bonferroni: $p < .001$), Desires (Bonferroni: $p = .001$) and Positive emotions (Bonferroni: $p < .001$). No significant differences emerged from the other comparison.

3.2. Performance at Th.o.m.a.s.: Gender differences

We evaluated possible gender differences in the performance at Th.o.m.a.s. with a repeated measure ANOVA with one within-subjects factor (type of scale, with four levels: scale A, B, C and D) and a between-subjects factor (type of group, with two levels: males and females) (see Fig. 4).

The analysis revealed an effect of scale type ($F_{(3,234)} = 21.41$; $p < .001$; $\eta^2 = .21$) and an effect of gender ($F_{(1,78)} = 5.18$; $p = .026$; $\eta^2 = .06$), finding that girls scored better than boys.

We investigated possible gender differences in the performance at each Th.o.m.a.s. subscale with a repeated measure ANOVA with one within-subjects factor (type of subscale, with three levels: Awareness, Relation, Realization) and a between-subjects factor (type of group, with two levels: males and females) (see Fig. 5).

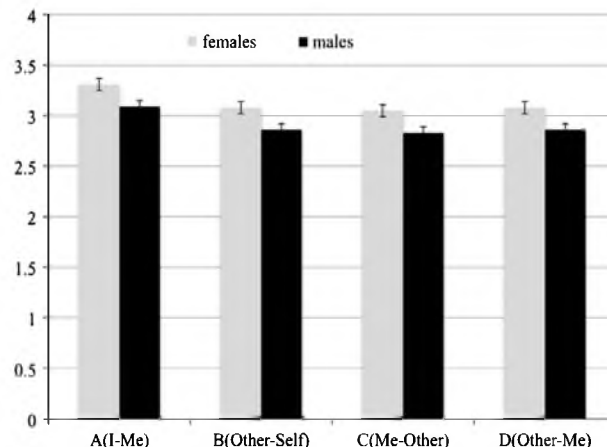
The analysis revealed an effect of subscale type ($F_{(2,156)} = 20.08$; $p < .001$; $\eta^2 = .2$) and an effect of gender ($F_{(1,78)} = 5.03$; $p = .028$; $\eta^2 = .06$), showing that females scored better than males in all the subscales.

We investigated possible gender differences in the performance at each Th.o.m.a.s. dimension with a repeated measure ANOVA with one within-subjects factor (type of dimension, with four levels: Beliefs, Desires, Positive emotions, Negative emotions) and a between-subjects factor (type of group, with two levels: males and females; see Fig. 6).

The analysis revealed an effect of dimension type ($F_{(3,234)} = 10.78$; $p < .001$; $\eta^2 = .12$) and an effect of gender ($F_{(1,78)} = 4.97$; $p = .029$; $\eta^2 = .06$). Females obtained higher scores than males.

3.3. Performance at classical ToM tasks, age and sex difference

To evaluate the performance of the participants at the other, more conventional ToM tests, we conducted a repeated measure ANOVA with one between-subjects factor (type of group, with four levels: 11, 13, 15, 17 years) and a within-subjects



Mean scores for each gender group at Th.o.m.a.s. scales, with standard error bars.

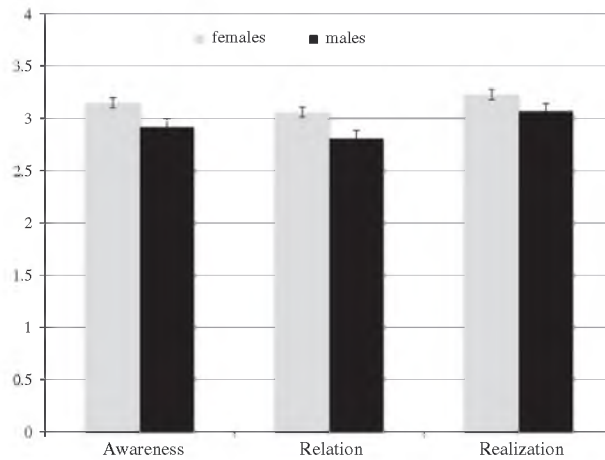


Fig. 5. Mean scores for each gender group at Th.o.m.a.s. subscales, with standard error bars.

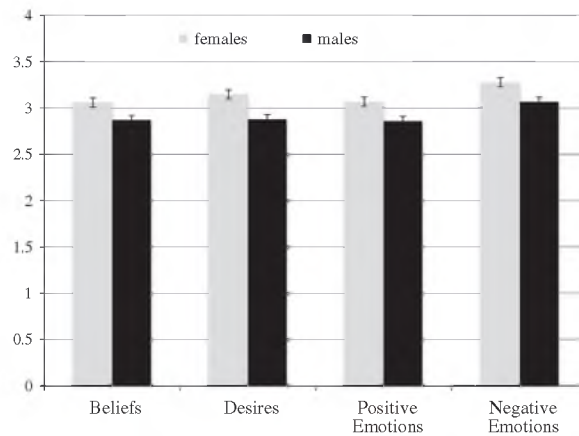


Fig. 6. Mean scores for each gender group at Th.o.m.a.s. dimensions, with standard error bars.

factor (type of test, with two levels: (Strange Stories and II order). Fig. 7 shows the mean scores for each age group at Strange Stories and the second-order tasks.

The analysis revealed an effect of type of test ($F_{(2,152)} = 14.94$; $p < .001$; $\eta^2 = .16$) but no effect of age group ($F_{(1,76)} = .26$; $p = .85$; $\eta^2 = .01$). *Post hoc* pairwise comparison revealed that the participants scored better at Strange Stories than at the second-order tasks (Bonferroni: $p < .001$).

We also evaluated possible gender differences in the performance at the standard ToM tasks. We conducted a *T* Test analysis that revealed no significant differences between the performance of males and females, both at Strange Stories task (*T* test: $t = .627$; $p = .53$) and at second-order task ($t = .374$; $p = .71$). The percentages of correct answers at Strange Stories task were 89% for the females and 87% for the males, and those at second-order tasks were 74% for the females and 76% for the males.

3.4. Performance at Th.o.m.a.s. scored by pragmatic assessment criteria

In principle, our results might be explained by an age-related improvement in ToM abilities, but also by an age-related improvement in communicative abilities. This is the reason why we also applied a set of communicative-pragmatic criteria to evaluate interviews.

We conducted an ANOVA analysis on the resulting scores which did not reveal any effect of the age on the adolescents' communicative performance at Th.o.m.a.s., scored by communicative-pragmatic criteria ($F_{(3,20)} = .133$; $p = .94$; $\eta^2 = .02$, see Fig. 8). A *post hoc* pairwise comparison revealed that there were no differences among in the adolescents' communicative-pragmatic performance, as they showed in answering at Th.o.m.a.s- belonging to the different age groups (Bonferroni: $p = 1$).

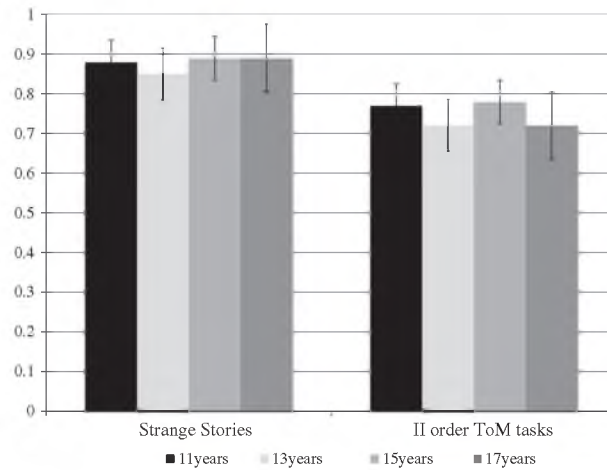


Fig. 7. Mean scores for each age group at ToM tasks, with standard error bars.

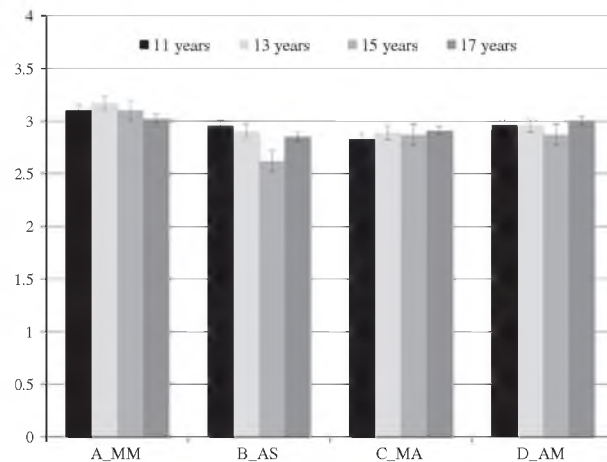


Fig. 8. Mean scores for each age group at the pragmatic assessment, with standard error bars.

4. Discussion

The aim of the research described in this paper was to provide an assessment as broad as possible of the development of Theory of Mind (ToM) abilities in preadolescence and adolescence.

In agreement with our hypotheses, the development of ToM does not end in childhood, but goes on at least through adolescence. The adolescents' performance improved with age in all the Th.o.m.a.s. scales, namely scale A, which investigates first-person ToM, scale B, which investigates first-order, third-person, allocentric ToM, scale C, which investigates first-order, third-person, egocentric ToM, and scale D, which investigates second-order ToM. However, the differences between age groups were not the same from each group age to the subsequent one: the effect of the age is consistent between 11 and 13 years, goes in the expected direction between 13 and 15 years, and then appears to stabilize. This datum seems in line with the first study of [Maylor et al. \(2002\)](#) to the effect that individuals aged between 16 and 29, when performing a set of advanced first-person ToM tasks including a subset of Strange stories ([Happé, 1994; Happé et al., 1999](#)) and new ToM stories constructed along similar lines, obtained a mean score of 4 against a maximum available of 7, thus showing no sign of a ceiling effect. In the light of the available literature, thus, it is unclear exactly when persons become able to fully manage their ToM ability (at least as measurable by the available experimental tasks).

As regards the distinction between first-person and third-person ToM, the participants performed better at scale A than at scale C. This was again in agreement with our expectations and may be related to the perception that a typical feature of preadolescence and adolescence is a tighter focus on the attempt to understand oneself than the others. Our results are also

in line with an fMRI study by Pfeifer et al. (2009) who found that the cerebral networks relevant to self perception are more active in adolescents than in adults while performing a direct self-reflection task.

However, whether humans are better reasoners in the first or in the third person or, in general, whether the notion of “self” is special, is a current matter of debate (see Gillihan & Farah, 2005). Our results support Goldman’s hypothesis (1993) to the effect that they can better reason about their own mental states than about those of the others, while other researchers argue in favor of the opposite view (e.g., Gopnik, 1993). Still in agreement with our hypotheses, the participants scored better at scale A, which assesses first-order ToM, than at scale D, which assesses second-order ToM. These findings are also in line with evidence gathered in children (Miller, 2009; Wellman & Liu, 2004) to the effect that the former kind of task is easier than the latter.

Again in accordance with our hypothesis, girls performed significantly better than boys at all the ToM components investigated through the scales, subscales and dimensions that make up the interview. However, the judges were not blind to the interviewee’s gender, which could be identified based on the syntax that he or she used.² Thus a possible explanation, which we are unable to rule out, is that the judges were influenced in their scoring by the stereotype that girls are more socially aware than boys. On the other hand, our datum is in line with other available evidence to the effect that females generally do have better mentalizing abilities than males (Baron-Cohen, Wheelwright, Hill, Raste, & Plumb, 2001) and perform better on social understanding tasks (Bosacki & Astington, 1999). It is generally reported in the literature that girls appear to be more socially competent, more compliant and more prosocial than boys; moreover, they also score higher in social perspective-taking and empathy tasks (Dodge & Feldman, 1990).

We had no hypotheses on a couple of issues, nor any hint could be found in the literature. One was the participants’ performance at scales B (Other-Self) vs. C (I-Other): both investigate third-person ToM, respectively from an allocentric and an egocentric viewpoint. We found no significant difference here.

The second was exploratory issue was possible differences between the three Th.o.m.a.s. subscales of Awareness, Relation, and Realization. The participants performed better at the Realization subscale than at the others. This suggests that adolescents are comparatively more capable of formulating and implementing strategies aimed at a desired state and of understanding what the desired states of the others are. Second best was the Awareness subscale, and last came Relation: adolescents thus appear to think that they are more capable of recognizing mental states *per se* than of understanding the causal links that they have with each other and with the behaviors that ensue.

Throughout the interview, Th.o.m.a.s. investigates different types of mental states (Beliefs, Desires, Positive emotions, Negative emotions). The participants scored higher at negative emotions than at the others. A possible explanation is that the many turbulent psychological and relational changes that characterize adolescence engender a sort of existential confusion which then leads individuals to reflect more deeply on their negative emotions.

We also administered a few more classical ToM tasks, namely two second-order tasks, namely the *Ice Cream Van Story* (Baron-Cohen, 1989) and the *Burglar Story* (Happé & Frith, 1994), and six stories taken from Happé’s *Strange stories* (Happé, 1994; Italian version by Mazzola & Camaioni, 2002). The evidence gathered was that first-order tasks are easier than second-order ones. This was in agreement with the results of the interviews, to the effect that the scores at scale A (first-order ToM) were better than those at scale D (second-order ToM), as well as with the literature available about children (Miller, 2009; Wellman & Liu, 2004). Our datum concerning adolescent’s performance at the Strange Stories appears to be in line with the second study reported in Maylor et al. (2002) showing that individuals aged between 18 and 27 years (mean age: 21) obtained a mean score of 7 against a maximum available of 8, that is, they did not reach ceiling; analogously, the 17-years olds that participated in our experiment obtained a maximum of 5 against a maximum available of 6.

By contrast, the performance both at the second order tasks and at Strange Stories revealed no significant age- or gender-related difference. In general, to our knowledge there is no evidence in the literature to the effect that the performance at these tasks should change in normally developing pre-adolescents and adolescents. While they are considered advanced ToM tasks, they were all originally developed for children (Baron-Cohen, 1989; Happé, 1994) and thus do not appear to be satisfactory tools for the investigation of mentalization at later ages.

Finally, we needed to check whether our results could be explained by an improvement of communicative-pragmatic abilities with age. To this effect we created a second set of criteria for scoring the interviews, this time based upon the criteria for the evaluation of such type of performance that could be found in the relevant literature (Bosco, Angeleri, Colle, Sacco, & Bara, 2013; Bosco et al., 2012; Penn, 1985; Prutting & Kitchner, 1987). We used these criteria to give a second score to 30% of the interviews, finding no significant age-related difference in the adolescents’ communicative-pragmatic performance, at least for the level of pragmatic ability required to answer Th.o.m.a.s.

These findings are in line with Nippold, Ward-Lonergan, and Fanning (2005), who found that, starting from 11 years of age, the grammatical ability of preadolescents is substantially close to that of adults, and with Lodge and Leach (1975), who found an acceleration in the understanding of figurative language and idiomatic meanings from 9 to 12 years of age, which testifies to the ability to handle complex meanings with the same level of ability as adults.

The substantially steady state reached by communicative-pragmatic abilities in adolescents is also testified by their inclusion, starting from 15 (Kim & Na, 2004) or 16 (Angeleri, Bosco, Gabbatore, Bara, & Sacco, 2012; Tombaugh, Kozak, & Rees, 1999) years of age, in normative data concerning the communicative performance of healthy adults. In conclusion, our data

² The interviews were conducted in Italian, a language that codes the feminine and the masculine genders in almost every noun, pronoun, adjective or verb.

confirm that ToM keeps maturing during adolescence, that this increase in performance is not explained by an increase of communicative-pragmatic abilities, and that ToM cannot be viewed as a unitary, all-or-nothing competence. Th.o.m.a.s. allows to investigate the changes that this ability undergoes over time, yielding, within a unitary framework, specific and comparable measures of different facets or components of ToM.

Acknowledgments

We wish to thank Francesca Condello, Alessandra Bassano, Tatiana Mantovan and Giovanna Alebardi for their assistance. This research was supported by Regione Piemonte, project "Institutions, Behaviour and Markets in Local and Global Settings (IINBEMA)".

Appendix A

A.1. The interview

This appendix contains the complete interview, divided into subscales. Th.o.m.a.s. being a semi-structured interview means that the interviewee's replies may sometimes anticipate questions that would have been the subject of a specific question at a later point. Analogously, explanations and examples may or may not be spontaneously offered by the interviewee. Therefore, a certain redundancy is present in the interview as it is reported here; this serves to remind the interviewer to ask for all the information needed, unless it has been spontaneously provided by the interviewee. The actual questions, however, are not as redundant and will vary slightly depending on what the interviewee may or may not have said spontaneously.

A.1.1. Scale A (I–Me)

[1] Do you happen to experience emotions that make you feel good? What? On what occasions? Can you give an example? (For example, in how you act or think, or in things that happen to you.)

[1a] (If the answer is negative) Why not?

[2] When you feel good, does that make any difference to you? What are the differences? Can you give an example of how you act or think, or of things that happen to you when you feel good?

[3] Do you happen to experience emotions that make you feel bad? What? On what occasions? Can you give an example?

[3a] (If the answer is negative) Have you ever asked yourself why?

[4] When you feel bad (or name negative emotions mentioned by the participant), does that make any difference to you? What are the differences? Can you give an example of how you act or think, or of things that happen to you when you feel bad?

[5] When you feel bad, do you feel you understand why? Can you give an example?

[6] Can you change your mood, when you want to? How? On what occasions? Can you give me an example? [6a] (If the answer is negative) Why not?

[7] Do you happen to have wishes, and know what you want? What? On what occasions? Can you give an example?

[7a] (If the answer is negative) Do you ever ask yourself why?

[8] Do you try to fulfill your wishes? How? On what occasions? Can you give an example?

[8a] (If the answer is negative) Why not?

[9] Do you succeed in getting what you want? How? On what occasions? Can you give an example?

[10] Can you explain why you succeed/do not succeed?

A.1.2. Scale B (Other–Self)

[11] Do the other persons happen to experience emotions that make them feel good? What? On what occasions? Can you give an example? (*The interviewee is left free to interpret the word "other" as she prefers. If the answer is too general, however, the interviewer would propose that a specific person be identified as the focus of reflection, like a friend, a familiar, or the spouse or partner.*)

[11a] (If the answer is negative) Why not, in your opinion?

[12] When the others feel good, does that make any difference to them? What differences does it make? Can you give an example of how they act or think, or of things happening to them when they feel good?

[13] And do the other persons happen to experience emotions that make them feel bad? What? On what occasions? Can you give an example? (*If the answer is too general and does not mention a specific "other" person, name a person mentioned by the participant earlier.*)

[13a] (If the answer is negative) Why not, in your opinion?

[14] When the others feel bad, does that make any difference to them? What differences does it make? Can you give an example of how they act or think, or of things happening to them when they feel bad?

- [15] In your opinion, when the others feel bad, do they understand why? Can you give an example?
 [15a] (If the answer is negative) Why do they not understand, in your opinion?
 [16] And, in your opinion, can the others change their mood when they want to? How? On what occasions? Can you give an example?
 [16a] (If the answer is negative) Why not, in your opinion?
 [17] Do the others happen to have desires and know what they want? What sorts of desires do they have? Can you give an example?
 [17a] (If the answer is negative) Why not, in your opinion?
 [18] Do the others try to fulfill their desires? How? On what occasions? Can you give an example?
 [18a] (If the answer is negative) Why do they not try, in your opinion?
 [19] In your opinion, do the others succeed in getting what they want? How? On what occasions? Can you give an example?
 [20] Why do they succeed/not succeed, in your opinion?

A.1.3. Scale C (I–Other)

- [21] Do you notice when the others feel good? When does that happen? Can you give an example?
 [21a] (If the answer is negative) Why do you not notice?
 [22] When you notice that another person feels good, does that make any difference to you? What differences does it make? Can you give an example, of how you act or think, or of the things that happen to you?
 [23] Do you notice when the others feel bad? When do you notice that? Can you give an example?
 [23a] (If the answer is negative) Why do you not notice?
 [24] When you notice that another person feels bad, does that make any difference to you? What differences does it make? Can you give an example of how you act or think, or of the things that happen to you?
 [25] When the others feel bad, do you understand why? Can you give an example?
 [25a] (If the answer is negative) Why can't you explain why other people feel bad?
 [26] Do you ever want to influence the mood of the others? How? On what occasions? Can you give an example?
 [27] Do you succeed in doing so? How? On what occasions? Can you give an example?
 [28] How do you explain the fact that you manage/do not manage to do so?
 [29] Do you think you understand the others' wishes? What sort of wishes do they have? Can you give an example?

A.1.4. Scale D (Other–Me)

- [31] Do the others notice when you feel good? When do they notice? Can you give an example?
 [31a] (If the answer is negative) Why do they not notice?
 [32] When the others notice that you feel good, does that make any difference to them? What difference does it make? Can you give an example of how they act or think when they notice that you feel good?
 [33] Do the others notice when you feel bad? When do they notice? Can you give an example?
 [33a] (If the answer is negative) Why do they not notice?
 [34] When the others notice that you feel bad, does that make any difference to them? What difference does it make? Can you give an example of how they act or think when they notice that you feel bad?
 [35] When you feel bad, do the others understand why? Can you give an example?
 [35a] (If the answer is negative) Why don't they understand?
 [37] Can the others influence your mood? How? On what occasions? Can you give an example?
 [38] How do you explain that they succeed/do not succeed in doing so?
 [39] Do you think that the others understand your desires? In your opinion, what sort of wishes do they think you have? Can you give an example?

Appendix B

B.1. Interview data coding grid

This appendix contains the grid for the coding and the insertion of the replies. Each question prompts the interviewee to supply a personal opinion regarding a specific aspect of how her mind or the mind of the others works. For example, question [3]: “Do you happen to experience emotions that make you feel bad?” investigates the interviewee's ability to identify her own positive emotions (Scale A: I–Me; subscale: Awareness; dimension: Positive emotion).

The scores for each question are inserted in the corresponding cell, to provide both qualitative and quantitative data for the various domains of theory of mind.

Scale:	A (I–Me)			B (Other–Self)		
Subscale:	Awareness	Relation	Realization	Awareness	Relation	Realization
Beliefs	×	5	10	×	15 (15a)	20
Desires	7 (7a)	8 (8a)	9	17 (17a)	18 (18a)	19
Positive emotions	1 (1a)	2	6 (6a)	11 (11a)	12	16 (16a)
Negative emotions	3 (3a)	4		13 (13a)	14	
Totals						

Scale:	C (I–Other)			D (Other–Me)		
Subscale:	Awareness	Relation	Realization	Awareness	Relation	Realization
Beliefs	×	25 (25a)	28	×	35 (35a)	38
Desires	29	26	×	39	×	×
Positive emotions	21 (21a)	22	27	31 (31a)	32	37
Negative emotions	23 (23a)	24		33 (33a)	34	
Totals						

Appendix C

C.1. Criteria for the communicative-pragmatic score

C.1.1. Score = 0

A score of 0 is attributed:

- when the interviewee shows low responsiveness or remains silent, however encouraged by the interviewer, thus violating the rules for turn-taking in conversations;
- when the answer is confused or disorganized, with frequent change of topic or “word salad”;
- when the answer is irrelevant to the topic of the question, or detached from the context;

C.1.2. Score = 1

A score of 1 is attributed:

- when the interviewee answers tangentially, without focusing on the topic of the question;
- when the interviewee limits herself to replying yes or no without adding further information, however encouraged by the interviewer; this is the case of too concise an answer, which does not take into account what the communicative partner has said and is characterized by poor informational content;
- when the answer is not well organized and lacks narrative coherence, specificity and accuracy;
- when an example is provided (spontaneously or after a request by the interviewer) which is not consistent with the topic of question or is not adequate to the context.

C.1.3. Score = 2

A score of 2 is assigned:

- when the subject, despite understanding the topic of the question, does not take into account the information provided by the communicative partner (the interviewer) and answers tangentially;
- when the interviewee fails to recover and/or edit his communicative intervention even after the interviewer’s suggestion that she may have misunderstood the perspective to take on the topic;
- when the answer is excessively repetitive (e.g., a tautological one) or concise, with no further explanation;
- when the interviewee uses prosodic and gestural elements which are not in line with the content expressed.

C.1.4. Score = 3

A score of 3 is assigned:

- to an answer which is not articulated or one which is coherent and correct, but with a narrative structure which is not fluent;
- to an answer which is coherent and consistent, but generic, stereotyped or only slightly contextualized;
- to an answer whose logic is good but not enough informative;
- to an answer which is consistent and relevant to the question asked, but includes prolonged pauses when the interviewee faces difficulties in taking her conversational turn;
- when the examples provided are not completely consistent with the topic of the question or are only partially adequate to the communicative context;
- when the interviewee uses prosodic and gestural elements inaccurately.

C.1.5. Score = 4

A score of 4 is attributed to a reply which:

- is coherent, detailed and organized, with significant, coherent and contextualized examples;
- is accompanied by prosodic elements which are correct, coherent and meaningful respect to the expressed content;
- apart from the sheer length of the answer, then, the judge should refer to particular pragmatic criteria such as turn-taking, topic management, the ability to take into account what has been said by the communicative partner, and the accuracy and the coherence of the answer.

References

- Abu-Akel, A. (2003). A neurobiological mapping of theory of mind. *Brain Research Reviews*, 43, 29–40. [10.1016/S0165-0173\(03\)00190-5](https://doi.org/10.1016/S0165-0173(03)00190-5).
- Adolphs, R. (1999). Social cognition and the human brain. *Trends in Cognitive Science*, 3, 469–479. [http://dx.doi.org/10.1016/S1364-6613\(99\)01399-6](https://doi.org/10.1016/S1364-6613(99)01399-6).
- Altman, D. G. (1991). *Practical statistics for medical research*. London: Chapman and Hall.
- Angeleri, R., Bosco, F. M., Gabbatore, I., Bara, B. G., & Sacco, K. (2012). Assessment battery for communication (ABaCo): Normative data. *Behavior Research Methods*, 44(3), 845–861. [http://dx.doi.org/10.3758/s13428-011-0174-9](https://doi.org/10.3758/s13428-011-0174-9).
- Aperly, I. A., Samson, D., & Humphreys, G. W. (2009). Studies of adults can inform accounts of theory of mind development. *Developmental Psychology*, 45, 190–201. [http://dx.doi.org/10.1037/a0014098](https://doi.org/10.1037/a0014098).
- Baron-Cohen, S. (1989). The autistic child's Theory of Mind: A case of specific developmental delay. *Journal of Child Psychology and Psychiatry*, 30, 285–297. [http://dx.doi.org/10.1111/j.1469-7610.1989.tb00241.x](https://doi.org/10.1111/j.1469-7610.1989.tb00241.x).
- Baron-Cohen, S. (1995). *Mindblindness: An essay on autism and theory of mind*. Boston: MIT Press/Bradford Books.
- Baron-Cohen, S., Wheelwright, S., Hill, J., Raste, Y., & Plumb, I. (2001). The 'Reading the Mind in the Eyes' Test Revised Version: A study with normal adults, and adults with Asperger syndrome or high-functioning autism. *Journal of Child Psychology and Psychiatry*, 42, 241–252. [http://dx.doi.org/10.1111/1469-7610.00715](https://doi.org/10.1111/1469-7610.00715).
- Blakemore, S. J. (2008). The social brain in adolescence. *Nature Reviews Neuroscience*, 9, 267–277. [http://dx.doi.org/10.1038/nrn2353](https://doi.org/10.1038/nrn2353).
- Blakemore, S. J., den Ouden, H., Choudhury, S., & Frith, C. (2007). Adolescent development of the neural circuitry for thinking about intentions. *Social Cognitive and Affective Neuroscience*, 2, 130–139. [http://dx.doi.org/10.1093/scan/nsm009](https://doi.org/10.1093/scan/nsm009).
- Bosacki, S. (2000). Theory of mind and self-concept in preadolescents: Links with gender and language. *Journal of Educational Psychology*, 92, 709–717. [http://dx.doi.org/10.1037/0022-0663.92.4.709](https://doi.org/10.1037/0022-0663.92.4.709).
- Bosacki, S. (2003). Psychological pragmatics in preadolescents: Sociomoral understanding, self-worth, and school behavior. *Journal of Youth and Adolescence*, 32, 141–155. [http://dx.doi.org/10.1023/A:31021861902463](https://doi.org/10.1023/A:31021861902463).
- Bosacki, S., & Astington, J. W. (1999). Theory of Mind in preadolescence: Relations between social understanding and social competence. *Social Development*, 8, 237–255. [http://dx.doi.org/10.1111/1467-9507.00093](https://doi.org/10.1111/1467-9507.00093).
- Bosco, F. M., Angeleri, R., Colle, L., Sacco, K., & Bara, B. G. (2013). Communicative abilities in children: An assessment through different phenomena and expressive mean. *Journal of Child Language*, 40, 741–778. [http://dx.doi.org/10.1017/S0305000913000081](https://doi.org/10.1017/S0305000913000081).
- Bosco, F. M., Angeleri, R., Zuffranieri, M., Bara, B. G., & Sacco, K. (2012). Assessment battery for communication: Development of two parallel forms. *Journal of Communication Disorders*, 45, 290–303. [http://dx.doi.org/10.1016/j.jcomdis.2012.03.002](https://doi.org/10.1016/j.jcomdis.2012.03.002).
- Bosco, F. M., Colle, L., De Fazio, S., Bono, A., Ruberti, S., & Tirassa, M. (2009a). Th.o.m.a.s.: An exploratory assessment of Theory of Mind in schizophrenic subjects. *Consciousness and Cognition*, 18, 306–319. [http://dx.doi.org/10.1016/j.concog.2008.06.006](https://doi.org/10.1016/j.concog.2008.06.006).
- Bosco, F. M., Colle, L., & Tirassa, M. (2009b). The complexity of Theory of Mind. *Consciousness and Cognition*, 18, 323–324. [http://dx.doi.org/10.1016/j.concog.2008.12.007](https://doi.org/10.1016/j.concog.2008.12.007).
- Bosco, F. M., Capozzi, F., Colle, L., Marostica, P., & Tirassa, M. (2014). *Theory of Mind deficit in subjects with alcohol use disorder: An analysis of mindreading processes*. [http://dx.doi.org/10.1093/alcac/agt148](https://doi.org/10.1093/alcac/agt148).
- Bosco, F. M., Friedman, O., & Leslie, A. (2006). Recognition of pretend and real actions in play by one- and two-year-olds: Early success and why they fail. *Cognitive Development*, 21, 3–10. [http://dx.doi.org/10.1016/j.cogdev.2005.09.006](https://doi.org/10.1016/j.cogdev.2005.09.006).
- Brothers, L. (1990). The social brain: A project for integrating primate behavior and neurophysiology in a new domain. *Concepts in Neuroscience*, 1, 27–51.
- Burnett, S., Sebastian, C., Coehen Kadosh, K., & Blakemore, S. J. (2011). The social brain in adolescence: Evidence from functional magnetic resonance imaging and behavioural studies. *Neuroscience and Biobehavioral Reviews*, 35, 1654–1664. [http://dx.doi.org/10.1016/j.neubiorev.2010.10.011](https://doi.org/10.1016/j.neubiorev.2010.10.011).
- Castellino, N., Bosco, F. M., Marshall, W. L., Marshall, L. E., & Veglia, F. (2011). Mindreading abilities in sexual offenders: An analysis of theory of mind processes. *Consciousness and Cognition*, 20, 1612–1624. [http://dx.doi.org/10.1016/j.concog.2011.08.011](https://doi.org/10.1016/j.concog.2011.08.011).
- Chiavarino, C., Bianchino, C., Brach-Prever, S., Riggi, C., Palumbo, L., Bara, B. G., et al. (2014). Theory of mind deficit in adult patients with congenital heart disease. *Journal of Health Psychology*, 2014. [http://dx.doi.org/10.1177/1359105313510337](https://doi.org/10.1177/1359105313510337).
- Choudhury, S., Blakemore, S. J., & Charman, T. (2006). Social cognitive development during adolescence. *Social Cognitive and Affective Neuroscience*, 1, 165–174. [http://dx.doi.org/10.1093/scan/nsi024](https://doi.org/10.1093/scan/nsi024).
- Coleman, J., & Hendry, L. B. (1999). *The nature of adolescence* (3rd ed.). New York: Routledge.
- Dodge, K., & Feldman, E. (1990). Issues in social cognition and sociometric status. In S. R. Asher & J. D. Coie (Eds.), *Peer rejection in childhood*. New York: Cambridge University Press.

- Dumontheil, I., Apperly, I. A., & Blakemore, S. J. (2010). Online usage of theory of mind continues to develop in late adolescence. *Developmental Science*, 13, 331–338. <http://dx.doi.org/10.1111/j.1467-7687.2009.00888.x>.
- Frith, C. (2007). The social brain? *Philosophical Transactions of Royal Society B*, 362, 671–678. <http://dx.doi.org/10.1098/rstb.2006.2003>.
- Frith, U., & De Vignemont, F. (2005). Egocentrism, allocentrism and Asperger syndrome. *Consciousness and Cognition*, 14, 719–738. <http://dx.doi.org/10.1016/j.concog.2005.04.006>.
- Frith, U., & Frith, C. (2010). The social brain: Allowing humans to boldly go where no other species has been. *Philosophical Transactions of Royal Society London B*, 12, 165–176. doi: 10.1098/rstb.2009.0160.
- Gillihan, S. J., & Farah, M. J. (2005). Is self special? A critical review of evidence from experimental psychology and cognitive neuroscience. *Psychological Bulletin*, 131, 76–97. <http://dx.doi.org/10.1037/0033-2909.131.1.76>.
- Goldman, A. I. (1993). The psychology of folk psychology. *Behavioral and Brain Sciences*, 16, 15–28.
- Goldstein, T., & Winner, E. (2012). Enhancing empathy and theory of mind. *Journal of Cognition and Development*, 13, 19–37. <http://dx.doi.org/10.1080/15248372.2011.573514>.
- Gopnik, A. (1993). How we know our minds: The illusion of first-person knowledge of intentionality. *Behavioral and Brain Sciences*, 16(1–14), 29–113.
- Happé, F. (1994). An advanced test of theory of mind: Understanding of story characters' thoughts and feelings by able autistic, mentally handicapped, and normal children and adults. *Journal of Autism and Developmental Disorders*, 24, 129–154. <http://dx.doi.org/10.1007/BF02172093>.
- Happé, F., Brownell, H., & Winner, E. (1999). Acquired 'theory of mind' impairments following stroke. *Cognition*, 70, 211–240. [http://dx.doi.org/10.1016/S0010-0277\(99\)00005-0](http://dx.doi.org/10.1016/S0010-0277(99)00005-0).
- Happé, F., & Frith, U. (1994). Theory of mind in autism. In E. Schopler & G. Mesibov (Eds.), *Learning and cognition in autism*. New York: Plenum Press.
- Hatcher, R., Hatcher, S., Berlin, M., Okla, K., & Richards, J. (1990). Psychological mindedness and abstract reasoning in late childhood and adolescence: An exploration using new instruments. *Journal of Youth and Adolescence*, 19, 307–326. <http://dx.doi.org/10.1007/BF01537075>.
- Kim, H., & Na, D. (2004). Normative data on the Korean version of the Western aphasia battery. *Journal of Clinical and Experimental Neuropsychology*, 26, 1011–1020. <http://dx.doi.org/10.1080/13803390490515397>.
- Laghi, F., Cotugno, A., Cecere, F., Sirolli, A., Palazzoni, D., & Bosco, F. M. (2014). *Bulimic symptoms. Theory of Mind and general psychopathology: A pilot study*. <http://dx.doi.org/10.1111/bjop.12054>.
- Lodge, D. N., & Leach, E. A. (1975). Children's acquisition of idioms in the English language. *Journal of Speech, Language and Hearing Research*, 18, 521.
- Maylor, E. A., Moulson, J. M., Muncer, A. M., & Taylor, L. A. (2002). Does performance on theory of mind tasks decline in old age? *British Journal of Psychology*, 93, 465–485. <http://dx.doi.org/10.1348/000712602761381358>.
- Mazzola, V., Camaioni, L. (2002). *Strane Storie: Versione italiana a cura di Mazzola e Camaioni*. Roma: Dipartimento di Psicologia dinamica e clinica, Università "La Sapienza".
- Miller, S. A. (2009). Children's understanding of second-order mental states. *Psychological Bulletin*, 135, 749–773. <http://dx.doi.org/10.1037/a0016854>.
- Moriguchi, Y., Ohnishi, T., Mori, T., Matsuda, H., & Komaki, G. (2007). Changes of brain activity in the neural substrates for theory of mind during childhood and adolescence. *Psychiatry and Clinical Neurosciences*, 61, 355–363. <http://dx.doi.org/10.1111/j.1440-1819.2007.01687.x>.
- Nichols, S., & Stich, S. P. (2003). *Mindreading. An integrated account of pretence, self-awareness, and understanding other minds*. Oxford: Clarendon Press.
- Nippold, M. A., Ward-Lonergan, J. M., & Fanning, J. L. (2005). Persuasive writing in children, adolescents, and adults: A study of syntactic, semantic, and pragmatic development. *Language, Speech, and Hearing Services in Schools*, 36, 125–138. [http://dx.doi.org/10.1044/0161-1461\(2005\)012](http://dx.doi.org/10.1044/0161-1461(2005)012).
- Onishi, K. H., & Baillargeon, R. (2005). Do 15-month-old infants understand false beliefs? *Science*, 308, 255–258. <http://dx.doi.org/10.1126/science.1107621>.
- Penn, C. (1985). The profile of communicative appropriateness: A clinical tool for the assessment of pragmatics. *The South African Journal of Communication Disorders*, 32, 18.
- Perner, J., & Wimmer, H. (1985). "John thinks that Mary thinks that...": Attribution of second-order beliefs by 5- to 10-year-old children. *Journal of Experimental Child Psychology*, 39, 437–471. [http://dx.doi.org/10.1016/0022-0965\(85\)90051-7](http://dx.doi.org/10.1016/0022-0965(85)90051-7).
- Pfeifer, J. H., Masten, C. L., Borofsky, L. A., Dapretto, M., Fuligni, A. J., & Lieberman, M. D. (2009). Neural correlates of direct and reflected self-appraisals in adolescents and adults: When social perspective-taking informs self-perception. *Child Development*, 80, 1016–1038. <http://dx.doi.org/10.1111/j.1467-8624.2009.01314.x>.
- Premack, D., & Woodruff, G. (1978). Does the chimpanzee have a theory of mind? *Behavioral and Brain Sciences*, 1(515–526), 1017. [10.1017/S0140525X00076512](https://doi.org/10.1017/S0140525X00076512).
- Prutting, C. A., & Kitchner, D. M. (1987). A clinical appraisal of the pragmatic aspects of language. *Journal of Speech and Hearing Disorders*, 52(2), 105.
- Sebastian, C., Viding, E., Williams, K. D., & Blakemore, S. J. (2010). Social brain development and the affective consequences of ostracism in adolescence. *Brain Cognition*, 72, 134–145. <http://dx.doi.org/10.1016/j.bandc.2009.06.008>.
- Shaw, D. J., Grosbras, M. H., Leonard, G., Pike, G. B., & Paus, T. (2012). Development of the action observation network during early adolescence: A longitudinal study. *Social Cognitive and Affective Neuroscience*, 7, 64–80. <http://dx.doi.org/10.1093/scan/nsq105>.
- Sowell, E. R., Peterson, B. S., Thompson, P. M., Welcome, S. E., Henkenius, A. L., & Toga, A. W. (2003). Mapping cortical change across the human life span. *Nature Neuroscience*, 6, 309–315. <http://dx.doi.org/10.1038/nn1008>.
- Taylor, S. J., Barker, L. A., Heavey, L., & McHale, S. (2012). The typical developmental trajectory of social and executive functions in late adolescence and early adulthood. *Developmental Psychology*. <http://dx.doi.org/10.1037/a0029871>.
- Tirassa, M. (1999). Communicative competence and the architecture of the mind/brain. *Brain and Language*, 68, 419–441. <http://dx.doi.org/10.1006/brln.1999.2121>.
- Tirassa, M., & Bosco, F. M. (2008). On the nature and role of intersubjectivity in human communication. *Emerging Communication: Studies in New Technologies and Practices in Communication*, 10, 81–95.
- Tirassa, M., Bosco, F. M., & Colle, L. (2006a). Sharedness and privateness in human early social life. *Cognitive Systems Research*, 7, 128–139. <http://dx.doi.org/10.1016/j.cogsys.2006.01.002>.
- Tirassa, M., Bosco, F. M., & Colle, L. (2006b). Rethinking the ontogeny of mindreading. *Consciousness and Cognition*, 15, 197–217. <http://dx.doi.org/10.1016/j.concog.2005.06.005>.
- Tombaugh, T. N., Kozak, J., & Rees, L. (1999). Normative data stratified by age and education for two measures of verbal fluency: FAS and animal naming. *Archives of Clinical Neuropsychology*, 14, 167–177. [http://dx.doi.org/10.1016/S0887-6177\(97\)00095-4](http://dx.doi.org/10.1016/S0887-6177(97)00095-4).
- Vogeley, K., Bussfeld, P., Newen, A., Herrmann, S., Happé, F., Falkai, P., et al (2001). Mind reading: Neural mechanisms of Theory of Mind and self-perspective. *Neuroimage*, 14, 170–181. <http://dx.doi.org/10.1006/nimg.2001.0789>.
- Vogeley, K., & Fink, G. R. (2003). Neural correlates of the first-person perspective. *Trends in Cognitive Sciences*, 7, 38–42. [http://dx.doi.org/10.1016/S1364-6613\(02\)00003-7](http://dx.doi.org/10.1016/S1364-6613(02)00003-7).
- Wellman, H. M., & Liu, D. (2004). Scaling of Theory-of-mind tasks. *Child Development*, 75, 523–541. <http://dx.doi.org/10.1111/j.1467-8624.2004.00691.x>.
- Wimmer, H., & Perner, J. (1983). Beliefs about beliefs: Representation and constraining function of wrong beliefs in young children's understanding of deception. *Cognition*, 13, 103–128. [http://dx.doi.org/10.1016/0010-0277\(83\)90004-5](http://dx.doi.org/10.1016/0010-0277(83)90004-5).