THEORY OF MIND AND EMOTION: STUDIES ON SCHOOL AGE CHILDREN

Coordinatore: Ch.ma Prof.ssa Antonella Marchetti
Tutor: Ch.ma Prof.ssa Olga Liverta Sempio

Tesi di Dottorato di: Giulia Cavalli
Matricola: 3380090

Anno Accademico 2006/2007
Contents

ACKNOWLEDGMENTS ................................................................................. 1

INTRODUCTION ......................................................................................... 2

CHAPTER 1. THEORY OF MIND: AN OVERVIEW ........................................ 5

1.1. What is Theory of Mind? .................................................................. 6
   1.1.1. Theories of ToM ...................................................................... 8
   1.1.2. What about emotions in ToM theories? ................................. 12
   1.1.3. ToM and neuroscience ........................................................ 13

1.2. Why to study Theory of Mind? ..................................................... 15
   1.2.1. Social functioning ................................................................. 16
   1.2.2. Emotional functioning ........................................................ 18

1.3. The development of mental states understanding ...................... 20
   1.3.1. Infants ................................................................................. 20
   1.3.2. Toddlers ............................................................................... 25
   1.3.3. School age children ........................................................... 27

1.4. How to study Theory of Mind? .................................................... 29

CHAPTER 2. THE VOICE TEST: A NEW ADVANCED THEORY OF MIND TASK FOR

ITALIAN SCHOOL AGE CHILDREN ......................................................... 32

2.1. Introduction: Why to create a new instrument? ......................... 36
   2.1.1. Theory of Mind tasks based on perceptive cues .................... 39

2.2. Phase 1: Construction of the Voice Test .................................... 50
   2.2.1. Preparation of the stimuli .................................................... 51
   2.2.2. Item reduction and administration procedure ..................... 53
   2.2.3. Further item reduction: the final version of the test .......... 57
   2.2.4. Conclusion ........................................................................ 63

2.3. Phase 2: Validation of the Voice Test ........................................ 65
   2.3.1. Method ............................................................................. 66
   2.3.2. Results ............................................................................. 69
   2.3.3. Discussion and conclusion ................................................ 74

2.4. Phase 3: Standardization of the Voice Test ............................... 75
FIGURES

Figure 2.1. Example of male eyes of the Eyes Test – child version ........................................44
Figure 2.2. Example of item of the emotion recognition in the face task ...............................48

CHARTS

Chart 2.1. Guilford’s difficulty indexes .................................................................60
Chart 2.2. The normal distribution of the Voice Test total score .......................................62
Chart 2.3. The Voice Test score increments with age ....................................................62
Chart 2.4. Age influence on the Voice Test performance ..............................................71
Chart 2.5. Correlation between the Voice Test and the Eyes Test ..................................72
Chart 2.6. Correlation between the 2nd order False Belief task and the Voice Test ..........73
Chart 2.7. Correlation between the Voice Test and the PPVT – R ..................................73
Chart 2.8. Voice Test total score distribution .............................................................78
Chart 2.9. Gender differences in the Voice Test .........................................................80
Chart 2.10. Age differences in the Voice Test ............................................................80
Chart 2.11. Correlation between the Voice Test and the PPVT – R .................................89
Chart 2.12. Correlation between the Voice Test and the RCPM ....................................89
Chart 3.1. Correlations between the Voice Test and the risk of depression .....................107
Chart 3.2. Correlations between the Voice Test and the risk of somatization .................108
Chart 3.3. Gender and age differences in the Voice Test performance ............................119

TABLES

Table 2.1. Guilford’s difficulty index for each item ....................................................60
Table 2.2. Total error score and total score of the Voice Test .......................................61
Table 2.3. Means and standard deviations ...............................................................70
Table 2.4. Correlations ...............................................................................................72
Table 2.5. Mean and standard deviation of the Voice Test total score in each group of age ..79
Table 2.6. Post hoc analysis .......................................................................................81
Table 2.7. Age 7 years (78-89 months): z-scores .....................................................82
Table 2.8. Age 8 years (90-101 months): z-scores ...................................................82
Table 2.9. Age 9 years (102-113 months): z-scores .................................................83
Table 2.10. Age 10 years (114-125 months): z-scores .............................................83
Table 2.11. Age 11 years (126-137 months): z-scores .............................................84
Table 2.12. Descriptive statistics of the Voice Test in the test and re-test conditions ........88
Table 2.13. Descriptive statistics in the test and re-test conditions for each group of age ……88

Table 3.1. Means and standard deviations .................................................................105

Table 3.2. Bivariate correlations between ToM and psychosocial risk .............................107

Table 3.3. Means and standard deviations .................................................................118
Abstract

Theory of Mind (ToM) is the ability to impute mental states – intentions, beliefs, thoughts, desires, emotions and other psychological states – to the self and to the others as a way of making sense and predicting behavior. Recently ToM researchers has been adopted a life span perspective – that leads to the construction of new instruments to assess ToM – and studied individual differences in ToM, including emotion understanding and the relationship between ToM and socio-emotional functioning. The present work is aimed to analyze deeper the development of mental states (both epistemic and emotional) understanding in children and its link with emotional difficulties, dealing with some novel topics within ToM studies, regarding ToM assessment in school age children with a new advanced task and the relationship between ToM and internalizing problems, characterized by emotional problems (while traditionally ToM researches study preschool children, using classical false belief tasks, and the link between ToM and externalizing behaviors). The first chapter reviews ToM studies: the main theoretical approaches; the link of ToM with the social and emotional functioning; the development of ToM; the methodological issue, concerning the assessment of epistemic and emotional states, in ToM studies. The second chapter presents the Voice Test, a new advanced ToM instrument, and its validation and standardization on an Italian school age children sample aged 6,5-11,4 years; the test assesses the ability to understand a wide range of complex mental states from vocal cues. Finally, in the third chapter it is studied the relationship between ToM and psychological risk in children and, in particular, it points out the link between poor ToM and frequent somatic complaints in a normal school age children population.
Acknowledgments

There are a number of people who I would like to thank for their guidance, encouragement and invaluable support. First I wish to express my sincerest gratitude to my PhD advisor, professor Olga Liverta Sempio, for everything she has taught me, for all her precious suggestions, her generosity in sharing her passion and knowledge in developmental research with me and for supporting all the time and encouraging me to achieve my potential. I am very grateful to professor Antonella Marchetti: she contributed to this project in numerous ways and I greatly appreciated her interest and suggestions on the design of this PhD project and her interesting comments on the topics of this thesis; I also thank her for supporting me throughout this work.

I would like to thank professor Rosa Angela Fabio: she was always helpful with the statistical analyses and research design and assisted me with every question I had.

I would like to thank my colleagues and friends, Annalisa, Barbara, Barbara and Eleonora, that providing moral support whenever needed, and all the members of the Theory of Mind Research Unit: they all making me this project an enjoyable experience.

Throughout my PhD years I have been continually helped by my family: their love and support has been essential. I have no words to thank my parents for all they do for me; they enable me to follow all my inclinations, in the various fields of my life.

I wish to thank Virgilio, for his love and presence, his precious suggestions and for sharing with me satisfactions and trouble of my work: he always encourage me to do my best.

Finally, I would like to thank all the children that participated in this project and their families, their teachers and school administrations, for their generosity and cooperation, making these PhD researches possible.
Introduction

The researches and reflections I present in this work, carried out during the three years of my PhD, are developed within the field of study of Theory of Mind (ToM). ToM is the ability to attribute mental states, such as beliefs, intentions, desires, pretending, knowledge, to oneself and others; it enables one to understand that mental states can be the cause of behavior and thus it can be used to explain and foresee actions (Premack & Woodruff, 1978).

For a long time ToM researchers focalized only on preschool children, on their comprehension of epistemic mental states (i.e., false belief) (Wimmer & Perner, 1983). Thanks to the studies conducted on this matter, it was possible to point out normative data and to create theoretical models on the development of mental state understanding during the preschool years (as I review in the first chapter).

In these last years researchers have subscribed to a broader definition of ToM (Bruner & Feldman, 1993; Hughes & Leekam, 2004), that includes a wide range of mental states, not only epistemic but also motivational and emotional states (Astington, 2001). Moreover recently researchers have adopted a life span perspective (Freeman, 2000; Khun, 2000), that leads for example to the construction of new instruments to assess ToM or its precursor from infancy to old age, even if it is hard to create ecological tasks that can assess the complexity of the mental states understanding.

Researchers are also interested in individual differences in ToM (Repacholi & Slaughter, 2003), not only in children with severe psychopathologies (i.e. autism spectrum disorders; Baron-Cohen, 1995), but also in typical children and children with various psychological problems. For example they studied this ability – understood also as a
mentalising style (Sharp, Coudace & Goodyer, 2007) – in children with externalizing behaviors (Liverta Sempio, 2002; Sharp, 2006). In fact ToM is supposed to be link to the social functioning (Moore & Frye, 1991; Astington, 2003), because people act on the basis of their mental representation of the world.

This thesis has the aim to analyze deeper and go forward in these new directions within ToM studies.

In particular it is focalized on typical school age children, that received not much attention, because of – I suppose – the difficulty to find tests that are able to assess the growing ability to understand complex epistemic and emotional mental states. In fact there are few advanced ToM instruments suitable for older children (Happè, 1994; Baron-Cohen et al., 2001b). Moreover these tasks use narrative or visual stimuli, that are very important in everyday life, but do not consider vocal cues that usually and often automatically we use during social relationship to grasp others’ mental states.

Finally, thinking to the Italian population, there are not ToM tests standardized for school age children.

In order to contribute to fill these methodological gaps it was created, validated and standardized a new advanced ToM task based on vocal stimuli for Italian school age children: the Voice Test. It assesses the understanding not only of complex epistemic mental states, but also of complex emotions, as explained in the second chapter.

Another new direction follows here, that contributes to study in depth ToM individual differences, regards the hypothesis that ToM is linked to the emotional functioning. Thus this work begins to investigate this matter, studying if children with emotional difficulties (i.e. internalizing problems, for example depression, anxiety and somatization) are impaired in mental states understanding. Because of the importance of the early mental health problems identification and consequently the early
psychological intervention, it can be important to study ToM development not only in children with proclaimed psychopathologies, but also in the normal children population, where some individuals could be at risk to develop psychological diseases. So the two researches presented in the third chapter investigated the link between ToM and emotional difficulties in a normal school age sample.

I conclude this brief introduction with the awareness that much work remain to be done, but also with the satisfaction to have contributed, theoretically, methodologically and empirically, to the investigation of ToM development and its link with emotional aspects and also to have begun to analyze the real-world consequences of children ToM (Astington, 2001) not only regarding the social functioning but also the emotional one, with attention paid to their understanding of both epistemic and emotional states.
Chapter 1

THEORY OF MIND: AN OVERVIEW

“An individual has a theory of mind if he imputes mental states to himself and others. A system of inferences of this kind is properly viewed as a theory because such states are not directly observable and can be used to make predictions about the behavior of others”, wrote in 1978 David Premack and Guy Woodruff (p. 516), two primatologists interesting in the possibility that chimpanzees are implicitly aware that different individuals can have different mental states, such as intentions and thoughts, and can use this ability to predict behavior.

From then on the concept of Theory of Mind (ToM) has been adopted by developmental psychologist (e.g. Bretherton, McNew & Beeghly-Smith, 1981; Wimmer & Perner, 1983; Baron-Cohen, Leslie & Frith, 1985; Perner & Wimmer, 1985) producing probably the fastest growing body of empirical research in psychology and giving rise to a wide range of different theoretical positions (Leudar, Costall & Francis, 2004). About thirty years after the research on children’s ToM began, this field of study continues to be a leading influence in the investigation of developmental psychology and psychopathology (Hughes & Leekam, 2004).

This chapter is aimed to answer to some questions about ToM. In particular, the chapter begins with the definition of ToM, that is strictly related to theoretical approaches assumed (what is ToM?); next comes a summary of the key reasons for the interest in ToM (why to study ToM?); then I synthesize the main research findings
about ToM (understood as a comprehension of both epistemic and emotional mental states) development from infancy to childhood (when ToM develops?); finally the chapter concludes with some consideration about the instruments used to assess this ability (how to study ToM).

1.1. What is Theory of Mind?

Theory of Mind is the ability to impute mental states to the self and to the others as a way of making sense and predicting behavior.

Over the years alternatives for the term “Theory of Mind” or “ToM” have also come into use (Whiten, 1994), such as “folk psychology” (Harris, 1992; Goldman, 1993; Stich & Ravenscroft, 1994), “intentional stance” (Dennet, 1987; Gergely et al., 1995; Griffin & Baron-Cohen, 2002), “mentalization” (Fonagy, 1991; Fonagy et al., 2002; Allen & Fonagy, 2006), “mentalizing” (Frith & Frith, 1999, 2003, 2006; Allen, 2003), “representation of subjectivity” (Battistelli, 1992, 1995) and “mindreading” (Baron-Cohen, 1994, 1995, 2005; Nichols & Stich, 2003; Coren, 2004; Tirassa, Bosco & Colle, 2006a). In psychology these terms are often used interchangeably (and the present work does so), but they have subtle differences.

Briefly, referring to Griffin and Baron-Cohen (2002), Frith and Frith (2006) and Sharp, Fonagy and Godyer (2006), the terms can be defined as follows, even if different theoretical approaches adopt quite different (i.e. restricted or wide) definitions (see the next section).
“Theory of Mind” or ToM (Premack & Woodruff, 1978) is the individual’s ability to represent themselves or others as having intentional and representational states (to desire, to believe, to know, to think...); this term refers specifically to the assumption that behavior is caused by mental states and often it is applied to the operationalization of the process we use to impute mental states.

“Folk psychology” (Harris, 1992; Goldman, 1993), that is called also belief-desire psychology or naïve/commonsense psychology, includes epistemic and emotional representational states and qualitative/phenomenal states, traits, disposition and empirical generalizations about behavior, too.

“Intentional stance” refers to how we predict behavior using intentional constructs and what intentional states really are; the term was introduce by Dennett (1987), who considered beliefs and desires as logical constructs not reducible to brain-states.

“Mentalization” (Fonagy, 1991) and “mentalizing” (Frith & Frith, 1999; Allen, 2003) refer to the process by which we make inferences about mental states (mentalistic insight); this process is supposed to be link to the mother-child attachment relationship and to the development of the self (with this meaning it is also adopted the term “reflective function”, used by Fonagy and Target, 1997, 2001; Fonagy et al., 2002).

The word “mindreading” (Baron-Cohen, 1995) is considered more theoretically neutral than other terms (Whiten, 1994), it has the same meaning of ToM, mentalization and mentalizing: it is the ability to understand that mental states, that are not observable, can be the cause of behaviors.

Finally, “representation of subjectivity” (Battistelli, 1992) underlines the role of subjective factors (e.g., epistemic and emotional states, moral judgment, meta-knowledge) and it could be referred to the different cognitive levels at stake when we understand humans’ actions.
The present work uses the terms “ToM”, “mentalization” and “mindreading” as synonymous, adopting a wide definition of this mental process, considered as a multicomponential ability (Bruner & Feldman, 1993).

Much ToM research has focused on a single children’s ability, that is the understanding of false belief. However, as Wellman and Liu (2004, p. 523) underlined, “many researchers believe that developing a Theory of Mind includes understanding multiple concepts acquired in an extended series of developmental accomplishments”. For example children can develop the understanding of desires, knowledge, beliefs, also those that are discrepant from reality (e.g., different desires, false beliefs) or across individuals (e.g., persons can have different beliefs about a situation); they also can develop the comprehension of emotions (real and apparent emotions) (e.g., Wellman & Liu, 2004); more sophisticated development are the understanding of metaphor, irony, double deceptions and complex narratives (e.g. Happè, 1994). Nevertheless, as Astington (2003) noted, even if researches follow the idea that ToM regards the comprehension of a wide range of mental states, they tend to focus on epistemic states, often omitting motivational states (intention and desire) and emotions\(^1\) (see also the paragraph 1.1.2.).

1.1.1. Theories of ToM

How do children come to understand themselves and other people as psychological beings? In other words, how they develop the ability to understand their own and

---

\(^1\) In the present work when I write “ToM” I always refer to the various mental states, including emotions, unless I specify that it is referred to epistemic mental states (e.g., false beliefs). When I write specifically about “emotions” I say if I refer to the scientific literature on emotions, out of the field of study of ToM.
others’ mental states? Several theories have been offered explanations to ToM development and nature.

The so called classical approaches are characterized by a piagetian (Piaget, 1975) conception of development (Astington, 1996), focused on individual factors rather than social ones. While contextualistic approaches are interested in the interpersonal relationships, underlining the interaction between cognition and social context.

*Classical approaches*

Classical ToM approaches are: theory-theory, simulation and modular theories (Astington, 1996).

Theory-theory approach (Wellman, 1990; Perner, 1991; Gopnik & Wellman, 1992; Gopnik & Meltzoff, 1997) suggests that humans have a theory used to reason about others’ mind. This theory, regarding persons’ mental states, is developed automatically and innately; it is like a scientific theory: mental states are theoretical notions that explain and predict behavior in the same way a scientist interprets data. So children develop a theory, a conceptual framework, and through it they can analyse other persons’ behaviors.

Simulation theory (Johnson, 1988; Harris, 1991, 1992) is based on the idea that children are aware of their own mental states through the introspection process and are able to simulate being in another’s shoes (a sort of role-taking), extrapolating from their own mental experiences. In other words, we understand others’ mental states not because we have a theory, but because we are able to simulate another person’s perspective from our own perspective.

Modular approach (Baron-Cohen, Leslie & Frith, 1985; Fodor, 1987; Leslie, 1987, 1994) postulates that ToM develops thanks to a neurological maturation of domain-
specific and modular mechanisms, that have an automatic functioning. These mechanisms are innate, although the experience is involved in their expression. For example, Baron-Cohen (1995; Baron-Cohen et al., 1995), according to Dennet’s (1987) intentional stance and to the evolutionary psychology, has adopted a phylogenetic perspective. He has hypothesized four modular mechanisms: the intentional detector (ID) and the eye direction detector (EDD), that developed in the first 9 months, regarding the ability to grasp volitional mental states of goal and desire and eye directions through visual, auditory and tactile perceptions; the shared attention mechanism (SAM) (9-18 months of life), that built triadic representations, through which the child can share attention with others on an external object; finally, the theory of mind mechanism (ToMM), activated around 3-4 years, that incorporated the previous mechanisms, using them in order to represent epistemic mental states.

**Contextualistic approaches**

The importance gave to interpersonal interactions for the psychological functioning (the “second wave” of studies; Bosacki & Astington, 1999) allowed ToM researchers to point out the role of the context in the mental states understanding development. For example, the socio-cultural approach (Bruner, 1990; Hobson, 1991; Feldman, 1992) theorized that ToM development is a social construction: the child is able to interpret himself and reality thanks to the cultural instruments and the narrative thinking (Bruner & Feldman, 1993). So the comprehension of how mental states guide self and others’

---

Leslie’s (1987) view is quite different from Baron-Cohen theorization; he hypothesized three modules: the theory of body mechanism (ToBy), developed around 3-4 months, regarding physical objects; the theory of mind mechanism (ToMM) regarding human mental states (8 months-2 years), and the selection processor (SP), developed around 4 years, that select information used by ToMM.
behaviors arises from the relationships, that develop in a specific cultural context (see also Olson & Astington, 1993; Astington, 1996).

The attention on interpersonal processes is underlined also by Fonagy and Target (2001), from a psychodynamic point of view, and by Meins, from a socio-constructivist perspective. In particular they focused on the mother-child attachment relationship (Bowlby, 1969, 1973, 1980). Fonagy and colleagues (Fonagy, Redfern & Charman, 1997; Fonagy & Target, 1997) suggested that mentalizing is a key determinant of children self-organization and found that the quality of attachment relationship, as well mothers’ reflective function (her ability to reflect on what is going on in her own and children’s mind), affect children ToM development⁴. Also Meins and colleagues (Meins, 1997; Meins et al., 1998, 2001, 2002, 2003) showed the caregivers’ important role for ToM development⁵. They found that in a secure attachment relationship, the mother is able to correctly identify child’s competences and to guide her child within the zone of proximal development⁶ (Vygotsky, 1978), moreover she is able to focus on her child’s mentalistic attributes (rather than physical or behavioral ones) when she describes her child (this ability is called “mindmindedness”; Meins et al., 1998)⁶.

---

³ Fonagy and colleagues (Fonagy, Redfern & Charman, 1997; Fonagy & Target, 2001) identified two different explanations about ToM development. One postulates that attachment relationship affects indirectly mentalizing development: the secure attachment allows children to engage in important social processes (symbolic play, language and social interaction with peers), that in its turn influences ToM. The second model hypothesizes a direct influence of mother-child attachment relationship on reflective function.

⁴ Meins (1997) theorizes that a secure attachment relationship influences ToM, through symbolic play and language.

⁵ The zone of proximal development is “the distance between the actual developmental level as determined by independent problem solving and the level of potential development as determined through problem solving under adult guidance, or in collaboration with more capable peers” (Vygotsky, 1978, p. 86).

⁶ About ToM socio-cultural, psychodynamic and socio-constructivist point of views see also Liverta Sempio & Marchetti, 2001a; Marchetti, 2002; Marchetti & Massaro, 2002; Lecciso, 2005.
Finally, Livertau Sempio and Marchetti (Livertau Sempio & Marchetti, 2001a, 2006; Livertau Sempio, 2002; Marchetti, 2002), from a socio-relational perspective, underlined that ToM, considered both a cognitive and affective ability, rises from a joint work between primary caregiver and infant, in which the dialogue between minds takes place through meaningful affective communication (i.e., the intersubjective exchange). So in this view it does not make sense to set beliefs against other mental states (such as emotions): all mental states are co-created within the affective relationships, that are socially and culturally situated.

The next paragraph deals with this question regarding how emotion understanding can or cannot be considered part of the mentalization ability.

### 1.1.2. What about emotions in ToM theories?

For a long time researches about ToM considered only epistemic mental states, especially false belief (Wimmer & Perner, 1983; for a critic: Bloom & German, 2000), but in these last years they subscribe to much broader definitions that encompass a wide range of mental states (Hughes & Leekam, 2004), including emotions. But what role have emotion in ToM theories? The answer obviously refers to “what ToM is” for each approach.

Briefly, following Liverta Sempio and Marchetti (2006), it can be noted that in the theory-theory approach (Gopnick & Wellman, 1992; Perner, 1991) emotions are considered as mental states among others; in the simulation approach, Harris (1989; 1998).

---

7 The socio-relational approach agreed with other perspectives, within the field both of psychoanalysis (e.g., Stern, 1985, 1998; Lyons-Ruth, 1998; Tronick, 1998; Sander, 2007) and the cognitive psychology (e.g., Liotti, 2005; Tirassa, Bosco & Colle, 2006a, 2006b), pointing out the importance of family relationships, without over- or underestimating infant’s abilities: intrapersonal and interpersonal processes are jointly responsible for ToM development.
1992) was one of the earliest researchers that linked ToM and emotions; while the modular approach referred only to epistemic mental states.

Inside cultural vision (Bruner, 1990; 1996; Bruner & Feldman, 1993) emotions could be considered integrated into the narrative frame: the narrative thinking, culturally deep-rooted, is on the basis of the “folk psychology”. In the vygotskian view of authors like Olson and Astington (1993; Astington, 1996), emotions are not considered or are merely part of contextual factors.

In Fonagy (2002; Fonagy et al., 2002) and Meins’ (Meins et al., 1998) perspectives, focused on the attachment relationship, ToM and emotions are linked inextricably. Also the socio-relational approach, as I explained at the end of the previous paragraph, considered emotions, as well the other mental state, growing within the affective relationships.

These different perspectives about the link between emotions and ToM influence the methodological choices, as I explain in the paragraph 1.4.

1.1.3. ToM and neuroscience

In this paragraph I summarize the recent findings regarding the neural basis of ToM; investigations on this matter only recently became feasible and in future they will be more and more sophisticated, giving us new data that could be used to build a wide and complex picture about ToM, its development, its link with other cognitive, emotional and social processes (e.g., language, executive function, metacognition, decision making, social and emotional functioning) and the links among the understanding of different mental states (such as emotional and epistemic).
A wide range of technique are used in order to study the neural structure of ToM, for example EEG, TMS, fMRI and psychophysiological techniques\(^8\) (Saxe & Baron-Cohen, 2006). Current researches (e.g., Fletcher et al., 1995; Gallagher et al., 2000; Adolphs, 2001; Stuss, Gallup & Alexander, 2001; Castelli et al., 2002; Walter et al., 2004; Frith & Frith, 2006; Saxe & Baron-Cohen, 2006; Singer, 2006) suggested the activation of the orbitofrontal cortex (with the superior temporal sulcus and the amygdala) and cortical frontal regions, in particular the medial frontal and prefrontal cortex, when persons are asked to infer mental states of others.

Results in the neuroscientific studies are not totally univocal, probably because they used different ToM tasks, consisting of narrative or perceptual stimuli, that require an implicit or an explicit mental state recognition, and investigated various mental states (depending on the definition of ToM adopted).

Recently a particular class of neurons, placed in the frontal cortex, called mirror neurons, discovered by Gallese, Fadiga, Fogassi and Rizzolatti (1996), researchers at the University of Parma, seems to be involved in imitation, intention understanding, empathy and language learning.

Mirror neurons are considered the neural correlate of mindreading (Gallese & Goldman, 1998; Gallese, 2003; 2006; Gallese, Keysers & Rizzolatti, 2004; Rizzolatti & Craighero, 2004; Rizzolatti & Sinigaglia, 2006). They are activated in relation to intentional behaviors performed both by self and by others, “providing a potential bridge between minds” (Williams et al., 2001). The damages in these structures could be responsible for ToM deficits of persons with autism, that show a defective “intentional attunement” (that is a direct form of experiential understanding of others), caused by a

\(^8\) EEG is the electroencephalography; TMS is the transcranial magnetic stimulation; fMRI is the functional magnetic resonance imaging (see Castelli, 2005); examples of psychophysiological techniques are the electrodermal activity and cardiovascular measures.
lack of embodied simulation, based on mirror neurons (Gallese, 2006; critical remarks: Hamilton, Brindley & Frith, 2007).

The discovery of this special neurons provides some evidence for the simulation theory, even if it constitutes an open issues (Borg, 2007). However it can be said that mirror neuron system seems to confirm the link between emotion and ToM, because it can be on the basis of cognitions and emotions within the relationships.

1.2. Why to study Theory of Mind?

Recent hypothesis suggest that ToM represents an evolved psychological capacity which evolutionary emergence occurred after the hominid-line split off; so it is most highly developed in humans than non-human primates (even if apes seem to be able to imitate, to self-recognize, maybe to deceive, to perspective-taking; Heyes, 1998; Tomasello, 1999; Call, 2007). As Brune and Brune-Cohrs (2006, p. 438) suggested, ToM “probably emerged as an adaptive response to increasingly complex primate social interaction”. So an answer to the question “why to study ToM?” deals with the social functioning, as also developmental researchers suggested (I explain this subject below, within this paragraph, and more in detail in chapter 3).

But adopting a wide definition of ToM, that includes also emotions, leads us to give another answer, that is interlaced with the first one, regarding the role of ToM in emotional functioning.
1.2.1. Social functioning

Since ToM is the ability to understand one’s own and others’ mental states, assuming they are on the basis of human behavior, it is evident its relevance for social and relational life: the understanding of and coordination with other peoples’ behavior is achieved through the connection between behavior and mind and thanks to this ability persons are able to give meaning to interpersonal behaviors. A good social adaptation (i.e., to be social competent) is possible, for example, if we can recognize epistemic and emotional mental states, especially those in contrast with reality (e.g., to distinguish appearance from reality, to realize the existence of diverse emotions, desires and beliefs, to understand false beliefs) (Moore & Frye, 1991; Baron-Cohen, 1995; Slaughter, Dennis & Pritchard, 2002; Astington, 2003; Cassidy et al., 2003).

Taking into account the mental states in order to interpret actions allows us to make sense of past behaviors, to affect present behaviors and to predict future behaviors. In this sense ToM has an important adaptive function. As Fonagy & Target (2001) suggested, children – giving meaning to social experiences – can create internal models or representations of the self and the others and can use these models to act appropriately during relationships.

The relationship between ToM and social functioning is complex and bidirectional: the development of ToM affects social competence and social relationships affects ToM development.

About the social implication of mindreading (usually operazionalized using the false belief task), it was found that a good ToM is related and predicts good relational competences (e.g., cooperative behaviors, positive relationships with peers) (Hughes & Leekam, 2004; Symons, 2004).
For example, children who are more able to understand others as intentional agents at nine months of age (see paragraph 1.3.1.), are more able to cooperate with peers at 24 months (Brownell, Ramani & Zervas, 2006); preschool children that understand the first order false belief have good communicative ability within peer relationships (Slomlowski & Dunn, 1996; Watson et al., 1999) and a good ToM seems also facilitate children’s transition to school and their ability to act in the school context (Astington & Pelletier, 1996; 2005). During the school-age period, children who can understand the first and second order false beliefs experienced more satisfaction for school life and are assessed as social competent by teachers (Astington, 2003); moreover ToM can help children to learn, not only through imitation and by instruction, but also collaboratively (Tomasello, Kruger & Ratner, 1993; Kruger & Tomasello, 1996; Tomasello, 2001). Nevertheless the role of ToM in the social competence development is not so clear (Liverta Sempio, 2002; Repacholi & Slaughter, 2003; see also paragraph 3.1.1.).

The researches on the role of social relationships in ToM development came most of all from studies regarding the mother-child relationship (Carpendale & Lewis, 2004), showing how this affective interaction can influence children’s understanding of mental states, as previously discussed from a theoretical point of view.

In particular, studies on family context underlined the role of some variables in children’s increasing mentalization ability, such as family structures (number of brothers/sisters) (Jenkins & Astington, 1996; Perner, Ruffman & Leekam, 1994; Arranz et al. 2002; Pears & Moses, 2003) and mother’s mental language and communicative style during mother-child interactions (Brown & Dunn, 1991; Meins et al., 2001; Hughes & Dunn, 2002; Ruffman, Slade & Crowe, 2002; Peterson & Slaughter, 2003; Symons, 2004; Taumoepeau & Ruffman, 2006; Lecce & Pagnin, 2007; review: de Rosnay & Hughes, 2006). Moreover, following Fonagy and colleagues (Fonagy,
Redfern & Charman, 1997; Fonagy & Target, 1997, 2001; Sharp, Fonagy, Goodyer, 2006; Fonagy, Gergely & Target, 2007) and Meins’s (1997; Meins et al., 2003) views, also the quality of the attachment relationship and the mother’s ability to mentalize influence the child’s ToM development.

1.2.2. Emotional functioning

ToM is important not only for the social functioning (Moore, Frye, 1991; Slaughter, Dennis, Pritchard, 2002; Astington, 2003; Cassidy et al., 2003), but also for the development of the self, in particular of the self-organization (Fonagy & Target, 1997) and self-awareness (Howlin, Baron-Cohen, Hadwin, 1999; Allen, Bleiberg & Haslam-Hopwood, 2003) (see also Lecciso, 2005): when children assign mental states to the self and to others, they also give meaning to psychological experience, first of all their own experience.

Allen, Bleiberg and Haslam-Hopwood (2003), adopting a psychoanalytic point of view, summarized some of the abilities connected to mentalizing, suggesting that it is crucial to our well being: mentalizing is the basis of a sense of identity, thanks to it we have a feeling of self-agency (that affects the experience of mastery and the sense of responsibility of our own behavior) and it allows us to have a sense of coherency and continuity; it is the basis of meaningful, sustaining relationships, in fact through the meeting of minds (Aron, 1996), that presupposes and affects mentalizing, we can feel connected to others and we can give and receive support (i.e. engagement in reciprocal relationships). Finally, they suggested that mentalizing is the key to self-regulation and it promotes the capacity to open emotional expression and sharing a full range of feelings.
Also Fonagy, Gergely, Jurist & Target (2002) suggested that mentalization has a regulation function. A self-regulation basic experience is the emotion regulation (Fonagy et al., 2002; Diamond & Aspinwell, 2003), that in general is the ability to think about one’s own emotions and to regulate them (Morris et al., 2007). They suggested (Fonagy et al., 2002; Jurist, 2005), connecting mentalization to the attachment theory frameworks (Bowlby, 1969, 1973, 1980), that the affect\(^9\) regulation arises from the attachment relationship, in which plays an important role the contingent and marked mirroring of child’s emotion (Gergely & Watson, 1996; Fonagy et al., 2002).

They coined the word “mentalized affectivity” to indicate the mature capacity for regulation of one’s own feelings and the capacity to discover the subjective meaning states of one’s own affects. It is a mature capacity for affect regulation, it includes the ability for appreciating difference between self and other and self-awareness.

While affect regulation, that is present from birth, is the ability to modulate emotions, mentalized affectivity not only modulates them, but also implies the comprehension of the impact of our representational world on affective experiences. Thus mentalized affectivity is a qualitatively different affect regulation and it can develop thanks to mentalization, that allows us to understand our and others’ mental states and to use this comprehension to self-reflect and regulate emotions.

\(^9\) In Fonagy and colleagues’ (2002) psychodynamic view the emotional experience within the mother-child attachment relationship is called “affect”. 
1.3. The development of mental states understanding

“The chronological order in which cognitive novelties emerge during childhood is a datum of central importance for the study of human cognitive growth”; following this Flavell’s (1972, p. 281) suggestion, in this paragraph I briefly review the main findings regarding ToM development, from infancy to childhood, trying to answer to the questions: what behaviors do children of different ages show that seem relevant to the development of knowledge about people?

1.3.1. Infants

Infants have many of the rudimentary skills needed for a more mature, fully functional ToM. They are precocious mentalization abilities, known as precursors of ToM, that show infants’ ability to grasp important aspects of the mind which guide their behaviors and influence others’ behaviors. Obviously, considering precursors implies a multi-componential view of ToM, beyond the false-belief comprehension, that comprises a wide range of inner states, not only cognitions, but also emotions and motivational states (intentions, desires) (Astington, 2001; Livertà Sempio & Marchetti, 2006).

The root of precursors is the babies’ sensitivity to social cues (e.g. eye contact and voice; especially those of mothers), that is present even hours after birth (Eisenberg, 1975; Haith, Bergman & Moore, 1977; De Casper & Fifer, 1980; Wolff, 1987; Slater & Butterworth, 1997; Bushnell, 2001); in the earliest phase of infancy children are also sensitive to contingent mother’s affective response, are able to discriminate her emotions and develop expectations about communicative interactions with her (e.g.,
Moreover, they are able to imitate, that is to convert an action plan originating from the other’s perspective into one’s own (Meltzoff & Moore, 1977, 1989, 1997, 1998; Legerstee, 1991; Kugiumutzakis, 1998; Meltzoff & Prinz, 2002; Nagy, 2006; Meltzoff, 2007) (see also Braten, 1998; Legerstee, 2005; Lavelli, 2007).

In the present brief review of the main ToM precursors, that develop during the first 24 months of life (reviews: Battistelli, 1997; Astington & Berriault, 2001; Flavell, 2004), I focus my attention particularly on those that are studied both in normal and clinical population (e.g. autistic children), often producing interesting debates not solved yet (Legerstee, 2005): joint attention, pointing, social referencing and symbolic play. They could be considered as cornerstones for socio-communicative development, pointing out that adaptive and interpersonal strategies are already present in infant’s mind. They grow in the parent-child affective relationship and they could be seen as intersubjective intentions, based on and aimed to build, moment by moment, intersubjective understanding of joint experience (Liverta Sempio & Marchetti, 2007).

During the first year children begin to learn how people differ from objects, understanding that animate agents are self-propelled and can move unpredictably (Legerstee, 1992, 2001; Woodward, 1998; Reddy, 2003). By around 9 months (but this ability seems to be already present at 5-7 months; Legerstee, Barba & DiAdamo, 2000; Legerstee & Barillas, 2003), they demonstrate the ability to understand intention, that is the understanding of other beings as intentional agents, that act in order to attain their aims (Phillips, Baron-Cohen & Rutter, 1992; Gergely et al., 1995; Meltzoff, 1995;

The ability to parse actions as intentional is evident when children begin to engage in joint attention, a triadic activity that involves the child, the adult and an external object/event toward which they both direct their attention: they have a shared goal and coordinate their actions for pursuing it. For example an infant can look at an object, in his line of vision (some months later also out of it, e.g. behind the child), that the mother is looking at: they achieve a common cognitive focus. During this period of life children can also use their understanding of intention when imitate others’ intentional actions, even if these actions don’t reach their goal (Bakeman & Adamson, 1984; Carpenter, Nagell & Tomasello, 1998; Phillips, Baron-Cohen & Rutter, 1998; Striano & Bertin, 2005; Tomasello et al., 2005; Bellagamba, Camaioni & Colonnese, 2006; Frischen, Bayliss & Tipper, 2007). These researches’ data testify the appearance in infants of a basic comprehension of others as intentional and mental agents, driven by internally represented goals (Legerstee, 2005).

Around the first birthday children actively establish joint attention through gesture, such as pointing (Bates, Camaioni & Volterra, 1975; Camaioni, Perucchini, Bellagamba & Colonnese, 2004; Liszkowski, 2005; Liszkowski, Carpenter & Tomasello, 2007).

Infants point in order to draw other’s attention to an object in the environment.

There are two modes of pointing: imperative and declarative. The first one is merely an instrumental action and appears generally around 10 months: the child requests an object or an action (e.g. to obtain an unreachable toy or to be picked up by the mother) and the adult is an instrumental agent (a means) who collaborates with him at a behavioral level. The declarative pointing is developed some months later (around 13 months) and serves to co-orient the child’s and the caregiver’s visual attention toward
the same object or event, so the dyad comes to share attention on something/somebody. The main difference between these two acts is that the imperative gestures are aimed to manipulate others’ behaviors or to manipulate the world through others’ behaviors (in this sense they don’t need a real mentalistic ability), whereas declarative ones seem to demonstrate that children are able to manipulate others’ minds or states of knowledge and imply an idea of others not as means but as subjects with a mind, that contains mental states (such as attention) that one can share and influence (Camaioni, 1993; Liszkwoski et al., 2004, 2006; Tomasello et al., 2005).

By the end of the first year of life, infants become also able to use the emotional expression of a familiar adult to regulate their behavior in unfamiliar and ambiguous situations (Sorce et al., 1985; Walden & Ogan, 1988; Moses et al., 2001; Stenberg, 2003; Walden & Kim, 2005; Striano, Vaish & Benigno, 2006). This ability, called social referencing, requires the previous acquisitions and the coordination of some important socio-communicative variables: the recognition and understanding of others’ emotional expression, the joint attention (sharing attention with adult on the novel situation) and, finally, the ability to modify one’s own behavior in response to the parent’s emotional expression (Desrochers et al., 1994; Moses et al., 2001; Carver & Vaccaro, 2007). For example, if a child is uncertain about how react (to approach or not) to an ambiguous stimulus (e.g. an unfamiliar toy or animal or a strange adult), he looks at his mother’s face and/or pays attention to his mother’s tone of voice (Vaish & Striano, 2004) not only to make sure that she is there, but also to use her expression in order to acquire information (i.e. to understand parent’s mental states) about the situation, to interpret it and to react appropriately to the stimulus.

Before the social referencing development, the infant is already able, around 9-10 months, to predict others’ behaviors from their emotions (Walden & Ogan, 1988);
then, during the second year, they begin to grasp others’ “emotional point of view”, for example they understand, from expressions and actions, if a person likes and desires an object that they doesn’t like (Denham, 1986; Wellman, Phillips & Rodriguez, 2000).

Another important change in infants mentalization ability occurs at about 18-24 months of age, when symbolic activities develop. In this period of life, children think not only about the world as it is, but also as it could be: they can imagine hypothetical situations, alternatives to reality (Leslie, 1987; Lewis & Ramsay, 2004). This ability is clearly evident in children’s pretend play, that consolidates during the third year. It is traditionally considered a meta-representational skill, because it requires the ability to transform symbolically objects and actions. For example children pretend when they use an object to represent a different absent object (e.g. a pencil could be used as a telephone, even if the child at the same time knows that it is really a pencil) or they treat an inanimate object as if it would be real (e.g. a child feed a doll, pretending it is a real baby).

In conclusion, precursors show a raising infants’ Theory of Mind, in fact these children’s behaviours seem to imply a primitive comprehension of one’s own and others’ minds and it has been observed that when this comprehension failed, such as in autistic children, later Theory of Mind development is impaired (e.g., Baron-Cohen, 1989, 1991, 1993, 1995, 2001).

All precursors involve relations with others: they grow and develop in the family relationship and, at the same time, influence it. Liverta Sempio and Marchetti (2007) suggested that they are a sort of “thinking with” the caregiver (a meeting of minds) and of “thinking about” one’s own and others’ mental states, built in child-caregiver relationship, even before the false belief understanding develops.
1.3.2. Toddlers

Preschool age is a developmental period in which there are many changes in mental states understanding, that progresses from an understanding of desire to an understanding of false belief (Wellman & Liu, 2004). During this period, from two years of age, the children’s ability to comprehend mental states is testified by the growing use of the mental state talk (i.e., children’s talk about inner states, using emotional, motivational and cognitive mental terms, such as “think”, “know”, “pretend”, “get angry”, “like”) (e.g., Bretherton & Beeghly, 1982; Dunn, Brown & Beardsall, 1991; Wellman & Bartsch, 1994; Dunn & Brown, 1993; Wellman et al., 1995; Hughes & Dunn, 1998; Ruffman, Slade & Crowe, 2002; Hughes, Lecce & Wilson, 2007; see also Lecce & Pagnin, 2007).

Around the end of the second years of life children understand that desires cause behaviors (Wellman & Bartsch, 1994), they can reasoning about desires and recognize that people can have different ones, for example they understand that if a person likes broccolis more than biscuits, when the person wants something to eat he will take broccolis and not biscuits, even if children’s desire is diverse (Repacholi & Gopnik, 1997). They also seem to grasp simple causal relations among desires and emotions, for example they understand that people will feel good if they get what they want and feel bad if they do not (Bartsch & Wellman, 1995; Flavell, 1999; this ability becomes more explicit around 3-5 years: Pons, Harris & de Rosnay, 2004).

During this early preschool period children can already realize that a person will see an object not only if the person’s eyes are aimed in the direction of the object, but also if there are not vision-blocking obstacles (Flavell et al., 1981; Flavell, 1988; Gopnik, Slaughter, Meltzoff, 1994), so they are able visual perspective-taking.
When children are three years old, they can understand the true belief, referred to reality (Wellman, 1991; Lohmann, Carpenter & Call, 2005), and engage in belief-desire reasoning (Wellman & Bartsch, 1994). They can also judge that they and someone else can have differing beliefs about the same situation (when the child does not know which belief is true and which is false) (Wellman & Liu, 2004).

Regarding the development of emotion understanding, Izard (1971) found that by the preschool years, most children can discriminate the facial expressions for happiness, sadness, anger, and fear. Thus they start to be able to recognize and name emotions on the basis of expressive cues and, around 3 years, they begin to understand how external causes affect the emotions of other children (Pons, Harris & de Rosnay, 2004).

At the age of four years, children are able to understand the false belief (Wellman, Cross & Watson, 2001; Flynn, 2006): they judge that a person can have a false belief about a situation, while at the same time they know that what they believe is true (they handle two different mental representation of reality). This ability is usually assessed with a verbal task, the first-order false belief task (Wimmer & Perner, 1983; Castelli, Lecciso & Pezzotta, 2001; Liverta Sempio et al., 2005; see also the paragraph 1.4.), but Onishit and Baillargeon (2005), using a non-verbal task, showed that 15-month-old infants can predict an actor’s behavior on the basis of her true or false belief about a toy’s hiding place.

Traditionally this ability is considered an important developmental stage, but – as I underlined before (paragraph 1.1.) – it represents only one aspect, among others, of people’s understanding of others’ mind, “nothing more, nothing less” and currently it is controversial how well standard false-belief tasks actually measure this concept (Bloom & German, 2000, p. B30).
At 4-5 years old, children also start to understand that a person’s beliefs (true or false) will determine his emotional reaction to the situation and are able to judge that a person can feel one emotion but display a different emotion (real-apparent emotion) (Harris, 1989; de Rosnay et al., 2004; Pons, Harris & de Rosnay, 2004; Wellman, Liu, 2004; Pons, Harris, 2005).

During the preschool period children develop not only the knowledge on the belief about physical objects (e.g., about an object’s identity or location, as operationalized in the false belief tasks), but also the knowledge about how to do things (such as to complete an homework). It requires not simply the understanding of perceptual access for recognizing how the knowledge is acquired (e.g., looking; Wimmer, Hogrefe & Perner, 1988; Pratt & Bryant, 1990), but also the understanding of other forms of knowledge sources, such as the activity of teaching (Ziv & Frye, 2004). Briefly, teaching implies two components: the awareness of a difference of knowledge (between teacher and learner) and the intentional activity (to increase others’ knowledge) (Olson & Bruner, 1996; Tomasello, Kruger & Ratner, 1993; Ziv & Frye, 2004). Researchers found that children’s teaching strategies (Wood et al., 1995; Astington & Pelletier, 1996; Ashley & Tomasello, 1998; Strauss, Ziv & Stein, 2002), as well learning ability (Tomasello, Kruger & Ratner, 1993; Kruger & Tomasello, 1996; Tomasello, 2001; Astington & Pelletier, 2005), are related to change in mental state understanding.

1.3.3. School age children

ToM development continues during the school age, but there are few researches on this period (studies are most of all on children with psychological difficulties or pathologies;
When children begin the primary school, they are able to interpret ambiguous situation using the reasoning on expectation (Pillow & Henrichon, 1996), to distinguish jokes from lies (Sullivan, Wimmer & Hopfield, 1995) and to understand the second order false belief (Perner & Wimmer, 1985), that requires making inferences about someone’s false attribution of belief. In the early school years children were found to understand the situational determinants of happy and scared and, some month later, also of sad and angry feelings (Harris et al., 1987) and become more able to predict emotions from beliefs (Pons & Harris, 2005).

Around 8 years, they understand mixed emotions or more sophisticated mental states, such as irony and metaphor (that imply the ability to go beyond the literal meaning of an utterance) (Ackerman, 1981; Marchetti, Massaro & Valle, 2007) and refer much more to mental states in their discourse (for example when they explain emotional reactions; Rieffe, Meerum Terwogt & Cowan, 2005), as well they better understand the polysemous meanings of the mental-state verb “know” (Booth & Hall, 1995).

In this period of life children’s developed ToM is also evident in their ability to lie consistently (Talwar, Gordon & Lee, 2007), to understand that a person may have multiple or even contradictory emotional responses to a situation and that morally reprehensible actions (e.g., to lie, to steal) cause negative feelings, while morally praiseworthy actions (e.g., to make a sacrifice, to resist a temptation) cause positive emotions (Pons, Harris & de Rosnay, 2004). So older children have knowledge of situations that evoke more complex emotions, such as pride, shame, guilt, gratitude, jealousy and worry (Harris et al., 1987).
As Pons, Harris and de Rosnay (2004) pointed out, children invoke different strategies for emotional control as they get older, for example at age 6-7 they refer mostly to behavioral strategies, while from 8 years old they start to acknowledge that psychological strategies can be more effective (e.g., denial, distraction). During the last years of primary school (9-11 years old) children can comprehend “faux pas” situation (Baron-Cohen et al., 1999), that requires the simultaneous mental representation of the perspective of the person who says something he should not have said, without grasping his mistake (the “faux pas”), and of the other person involved that can be hurt/irritated.

The growing mental states understanding is assessed with advanced tasks (described in chapter 2), focused for example on metaphor, irony and double deception (e.g., Happè, 1994), that are supposed to be acquired later in development than the understanding of desire, emotion, knowledge and belief (Brune & Brune-Cohrs, 2006).

1.4. How to study Theory of Mind?

In this paragraph I introduce the topic of ToM assessment, that is in-depth analyzed in the second chapter.

The classical instrument used to assess the first order false belief understanding is the “Sally-Ann task” (Wimmer & Perner, 1983; modified versions: Baron-Cohen, Leslie & Frith, 1985; Siegal & Beattie, 1991; Italian version: Liverta Sempio & Marchetti, 2001b). The child is shown a scenario illustrated, which can be enacted by puppets or real people.
There are two characters: Sally and Ann; Sally has a basket; Anne has a box. Sally puts a marble in her basket and then leaves the scene. While Sally is away and cannot watch, Anne takes the marble out of Sally's basket and puts it into her box. Sally then returns and the child is asked “Where will Sally look for her ball?”. To answer this question the child must realize that Sally has not seen the marble being moved and, therefore, that Sally has the false belief that the marble is still in the basket.

As I explained in the previous paragraph, normal children pass this task around the age of four. It seems that the difficulties of younger preschool children to understand the false belief in this task are not due to problems in the mnestic or language domain (Lewis & Osborne, 1990; Perner, Leekam & Wimmer, 1987; Wellman, Cross & Watson, 2001), even if language is strongly related to ToM development (Malle, 2002; Antonietti, Liverta Sempio & Marchetti, 2006; Milligan, Astington & Dack, 2007).

Another first order false belief task is “The deceptive box” (Perner, Leekam & Wimmer, 1987; Italian adaptation: Liverta Sempio & Marchetti, 2001c), that assesses the understanding of both one’s own and other’s false belief. This task, as the second order false belief task “Look prediction” (Sullivan, Zaitchik & Tager-Flusberg, 1994; Italian adaptation: Antonietti et al., 1999), are illustrated in the second chapter.

Most of ToM studies used these classical tests, but they are focused only on the false belief epistemic mental state. Other tasks assess basic or mixed emotions understanding (e.g., Steele et al., 1999; Ruffman, Slade & Crowe, 2002), some other studies both beliefs and emotions together (e.g., Symons et al., 1997; De Rosnay & Harris, 2002) or separately (Cutting, Dunn, 1999; Repacholi, Trapolini, 2004; Pears, Fisher, 2005). In the next chapter I review tasks that assess complex epistemic and emotional mental states (e.g., Happè, 1994; Baron-Cohen et al., 2001a, 2001b; Golan et al., 2007a).
This variety of tasks reflects the different ToM theories, described in the first paragraph of this chapter, that operationalized their definition of mindreading in different ways, performing different methodological choices. As Liverta Sempio and Marchetti (2006) showed, the more ToM definition is individualistic and decontextualized, the more researches have a correlational nature and use instruments referred to single areas (for example, false belief task and emotion recognition task separately). Vice versa the more ToM definition links cognition and emotions, the more researches use instruments that point out this connection between different aspects of mental functioning (e.g., Fonagy & Target, 1997). Thus some researchers, as I have showed before, include emotions into the set of mental states that constitutes ToM, some others conceived ToM quite exclusively as false belief understanding, putting emotions “outside” the set of mental states studied by the ToM paradigm.

A future promising direction of research should try to propose a more emotionally charged view of theory of mind, including all basic emotions, as well as mixed and complex emotions, within its objects of investigation.

Moreover it can be noted that most of ToM instruments assess the understanding of others’ mental states, not considering one’s own mental stated comprehension. Future researches should investigate also this matter and it could be interesting to better understand the relationship between one’s own and others’ epistemic and emotional states understanding, also in order to throw light upon children social and emotional functioning, that could be linked with the ability to understand mental states both of others and oneself.
Chapter 2

THE VOICE TEST: A NEW ADVANCED THEORY OF MIND

TASK FOR ITALIAN SCHOOL AGE CHILDREN

As I reviewed in the first chapter, researchers traditionally have been considered Theory of Mind (ToM) as a false belief understanding (Wimmer & Perner, 1983; Perner & Wimmer, 1985), assessed usually in preschool children.

In the present work I adopt a wide definition of ToM (Bruner & Feldman, 1993; Hughes & Leekam, 2004), implying a multi-componential view of it (Wellman & Liu, 2004), beyond the false-belief comprehension, that comprises a wide range of inner states, not only cognitions, but also emotions and motivational states (Astington, 2001). This wide point of view first of all allows ToM researchers to investigate ToM in infants and in older children, before and after the false belief understanding develops. I don’t want underestimate nor overestimate the importance of the false belief comprehension, that certainly is an unequivocal marker of mentalistic understanding (Wellman, Cross & Watson, 2001), but I consider it just as an aspect, among others, implied in ToM.

Secondly, the wide definition adopted here brings researchers to find ecological instrument, nearer to real life than classical false belief tasks, in order to assess on the one hand ToM precursors (see paragraph 1.3.1.) in infants and on the other hand the growing mentalizing ability, that become more and more complex in school age
children. This means to use in ToM assessment the same communication “channels” we use in everyday life when we recognize mental states, using them to make sense and predict behaviors, that are the visual (i.e., what we see, such as actions and facial expressions) and auditory (i.e., what we heard) channels, besides the “cognitive” channel (it includes, for example, the memory ability, the comprehension and production of language, the executive function; e.g., we use this channel when we interpret some previous facts as cause of the present behavior and use them to understand the mental state below or when we give meaning to the words that a person says in order to understand his mental states).

For example, when we see a man frowns his brows, gesticulates animatedly (visual channel) and shouts (auditory channel) we can understand that he is probably angry. Or when we heard a woman says to a person, with a faltering voice (auditory channel), giving quick looks around (visual channel), in the middle of a busy crossing: “But I must turn right or left?” (cognitive channel), we can give meaning to her behaviors thinking that she is puzzled. Social functioning, therefore, requires the rapid processing of various stimuli conveyed principally by facial expression, vocal prosody and expression in the eyes.

Usually we use many information regarding behaviors (such as actions, verbal and non verbal behavior), that we grasp with all our senses (mainly sight and hearing), to impute mental states to the self and to the others. It is not simple to operazionalize what we experience daily when we mentalize; a solution could be to consider one aspect at a time. Researchers were focused particularly on aspects referring to what I called visual and cognitive channel. For example, some researches studied if children are able to predict, from the knowledge of the antecedent (e.g., a boy receives an unwelcome Christmas present from his parents), which mental state a story character has (e.g., he
will be sad or angry, or he will mask his real emotion and will say a lie to his parents because he doesn’t want to offend them) (e.g., Happè, 1994; Baron-Cohen et al., 1999; Muris et al., 1999; Pons & Harris, 2000). Some other researches analyzed if children are able to understand mental states from visual cues, such as persons’ facial expressions or schematized faces (e.g., Baron-Cohen, Wheelwright & Jolliffe, 1997; Rutherford, Baron-Cohen & Wheelwright, 2002; Golan, Baron-Cohen & Hill, 2006; Golan, Baron-Cohen, Hill & Golan, 2006).

Whereas in the field of study of ToM there are few researches focalised on the mental state understanding through the auditory channel. This is quite surprising considering that for example from early months of life infants are sensitive to vocal cues (Mastropieri & Turkewitz, 1999), respond appropriately to communication and they are able to grasp and differentiate emotions only hearing vocal cues (Caron, Caron & MacLean, 1988; Soken & Pick, 1999). Recently Vaish and Striano (2004), studying the social referencing – a ToM precursor (see paragraph 1.3.1.) – found that voice, even without a visual reference, is more potent than facial cues in guiding infants’ behavior. Perception researches have shown that human beings are able to identify emotions from verbal intonation differences in speech (Bachorowski & Owren, 1995; Banse & Scherer, 1996) and that listeners find auditory cues (such as tone, pauses, hesitations) more informative (about a speaker’s intention to lie) than visual information from the face (Wiseman, 1995).

For example adults rely to paralinguistic features (intonation) of the verbalization to label the speaker’s emotional state, when he presents contradictory emotional information (e.g., happy situation is described with sad paralanguage), like older children, that increase this ability from 5 to 10 years old; on the contrary younger children (4 years old) rely mostly to the linguistic content (Morton & Trehub, 2001).
The auditory channel is therefore apt to be especially effective in emotional communication (e.g., Sorce, Emde, Campos & Klinnert, 1985; Baldwin & Moses, 1996); linguistic and paralinguistic features of emotional speech are processed separately in the brain: verbal content is processed in the left hemisphere, while intonation is processed in the right one (McNeely & Parlow, 2001; Wildgruber et al., 2005).

The sensitivity to cues is a highly adaptive skill and, as suggested Vaish and Striano (2004, p. 266), and it deserves much more attention in the ontogeny of human social cognition. I could add that it deserves also attention in the following developments of mentalization, in fact there are not ToM researches that use this specific channel in older children, from the typical population.

Only some studies in adults with and without autism spectrum conditions, conducted mainly by Baron-Cohen and his team (see the descriptions in the next paragraph), used vocal cues, alone or with facial expressions, to assess ToM and to create training programmes to improve the mind-reading ability.

In this chapter I present a new ToM instrument, called Voice Test, that assesses children ability to recognize a wide range of mental states from vocal cues. First I describe why and how the Voice Test was created; then the works of validation and standardization of the instrument on a typical Italian school age children population; finally I suggest some possible uses of it and future directions about the assessment of ToM in children.
2.1. Introduction: Why to create a new instrument?

ToM is a complex, developmental phenomenon, which implies certainly more than just the understanding of false belief. There is a need for assessment tools that measure the developmental progression of ToM in a broader age range. Nevertheless a paucity of studies has explored the development of mentalization ability in school age children belonging to a normal population. Referring to the emotion development, a similar criticism has recently raised also within the field of study of emotion understanding (e.g., Pons, Harris, 2005) and emotion recognition from nonverbal cues (e.g., Rothman & Nowicki, 2004; Tonks et al., 2007). I guess because on the one hand (theoretical point of view) such ability has traditionally been thought to develop in early childhood and remain relatively stable throughout the rest of childhood and adult life (Battacchi, 2004; Nadel & Muir, 2005). On the other hand (methodological point of view) ToM assessment become more difficult in older children and it is hard to find instruments that can investigate the complexity implied in the everyday mental states understanding and in its link to other variables (such as other cognitive and emotional skills and personality aspects).

In these last years there is a increasing trend to find new instruments, modifying tasks already existent or creating new ones, able to assess ToM both in a normal and clinical population. These tasks are called “advanced” or “high-level” tasks (e.g., Happè, 1994; Baron-Cohen et al., 1997; Jolliffe & Baron-Cohen, 1999; Kleinman, Marciano & Ault, 2001; Brent et al., 2004), because they assess the recognition of complex mental states, that classical tests cannot grasp.
There are two main typologies of ToM advanced tasks (Liverta Sempio et al., 2006): tests based on perceptive cues and narrative tasks. This last typology of test assesses the ability to infer epistemic (e.g., irony, white lie, persuasion) and emotional (e.g., moral and mixed emotion) mental states from short stories, usually presented with pictures that facilitate children’s comprehension of the plot. Narrative tasks are, for example, the “Strange Stories” by Happè (1994; Italian translation: Mazzola & Camaioni, 2002), the “Faux Pas” by Baron-Cohen, O’Riordan, Stone, Jones and PLAisted (1999; Italian translation: Liverta Sempio, Marchetti & Lecciso, 2005b), the “TOM Test” by Muris and colleagues (1999) and the “Test of Emotion Comprehension (TEC)” by Pons and Harris (2000; Italian standardization: Albanese & Molina, in press).

Briefly, the “Strange Stories” test consists of 24 short stories about everyday social communication abilities, implying a mentalistic interpretation of behaviors (such as pretence, joke, lie, white lie, figure of speech, misunderstanding, persuasion, sarcasm, double bluff, contrary emotions, appearance/reality and forgetting); children are asked to explain why the story’s character acts like that; they usually pass this test around 8-9 years old.

The “Faux Pas” consists of 10 stories in which a character has a socially inappropriate verbal behavior (called faux pas); children judge if someone in the story said something he/she shouldn’t have said, because it would be hurtful; they correctly understand faux pas around 11 years old. The “TOM Test”, used with children between 5 and 12 years of age, comprises 78 questions about stories’ characters, subdivided in three subscales of growing difficulty: precursors of ToM (e.g., recognition of emotions, pretence), first manifestations of what the Authors called a “real ToM” (e.g., first order false belief understanding) and more advanced aspects of ToM (e.g., second order false belief understanding, comprehension of humour).
Finally, the “TEC” assesses, in 3-11 years old children, specific components of the understanding of emotion, hierarchically organised in three groups in function of their level of difficulty: the first group of components focuses on external aspects of emotions (e.g., recognition of facial expressions, understanding the impact of situational causes on emotions); the second one regards various mental aspects of emotions (e.g., the understanding of the role of beliefs and desires on emotions); and the third group assesses children’s understanding of the way in which an individual can think about an emotionally charged event from more than one perspective (e.g., understanding of mixed feelings, cognitive control strategies).

The tests I mentioned up to now imply the use of the channel I called before the cognitive one, in fact characters’ mental states are inferred only from events understanding (antecedents, behaviors and statements, described in the story). Children have not information about characters’ facial expressions or tone of voice, so these tasks do not require to refer to the perceptive channels to infer mental states. In our daily relational life it is very unlikely a lack of perceptive cues, moreover visual (such as facial expression, gaze, gestures, postures) and auditory (such as tone and rhythm of voice) cues are extraordinary informative signals of mental states. These cues are parts of human nonverbal behaviors (e.g., Darwin, 1872; Ekman & Friesen, 1969; Hinde, 1979) and their codification is important for social relationships (Ricci Bitti & Zani, 2002), in order to give meaning to one’s own and others’ behaviors.

In the field of ToM studies, in despite of the big amount of data analyzed by general psychology about the nonverbal behavior, only recently the attention has been turned to the ability to infer mental states also from perceptive cues. In the next paragraph I describe various tasks in which visual and/or vocal signals give information about persons’ epistemic or emotional mental states.
2.1.1. Theory of Mind tasks based on perceptive cues

This presentation of tasks based on perceptive cues created in the field of ToM follows a chronological order, because it suggests how ToM studies develops, both in theoretical and in methodological background, and it can better explain why there was a need to create a new instrument with vocal cues.

Visual Theory of Mind tasks

One of the first study that used a task based on perceptive cues was developed in 1992 by Baron-Cohen and Cross; using photographs of persons’ facial expression, they found that normally developing 4 years old children can infer when someone is thinking from the direction of gaze. In order to analyze deeper this result, Baron-Cohen, Riviere, Fukushima, French, Hadwin, Cross, Bryant and Sotillo (1996) investigated how cognitive mental states (and not only the six basic emotions: happiness, sadness, anger, disgust, fear and surprise\(^{10}\)) can be inferred from facial expressions.

At the time of their study, they could refer to a lot of works about emotion recognition (e.g., Ekman & Friesen, 1969; Izard, 1971; Zajonc, 1985), while the recognition from face of cognitive mental states were less studied. Only the cognitive states of deception (Ekman & Friesen, 1975; Kleinke, 1986) and interest (Izard, 1971; Ekman & Oster, 1987) focalized some researchers attentions, even within the field of study of emotions.

\(^{10}\) The authors (Baron-Cohen et al., 1996) agreed with Ekman’s (1992, 1999) conception of six basic emotions. But theorists disagree about how many such emotions there are (generally between six and ten) and which emotions are basic (emotion sets include fear, anger, joy, sadness, and disgust; some include surprise, shame, interest, etc.) (e.g., Izard, 1971; Frijda, 1986; see also: Ortony & Turner, 1990; Anolli, 2002; Solomon, 2002), so in the next pages I mention other studies that included a different number of basic emotions.
The aim of Baron-Cohen and colleagues’ (1996) study was to investigate whether adults (from a range of social classes, age groups and occupations) are able to perceive a range of mental states and whether this ability is stable across differing cultures. So the tested individuals from European (Britain, Spain) and non-European (Japan) cultures using illustrations of faces took from painting of famous painters; subjects were asked to choose, between two mental words, which best described what the person in the picture was feeling or thinking. They tested 11 different facial expressions (wary, astonished, contempt, recognize, threaten, regret, worried, distrust, revenge, guilt and scheme).

They found that adults have aptitude for reading a wide range of mental states in the face, in a cross-culturally similar way (with the exceptions of wary, guilt and scheme). Finally, Baron-Cohen and colleagues (1996) analyzed if and when children, from 8 to 11 years old, develop the ability to recognize the same set of mental states in the illustrations used in the previous experiments. They did not find age differences and concluded that by the age of 8 children are already very competent at reading the mind in the face, beyond the narrow set of basic emotions.

In the following years Baron-Cohen with his colleagues, within the Autism Research Centre (University of Cambridge), created new instruments based on perceptive cues, searching more subtle tests of mindreading especially to assess ToM in adult with autism, because he found that false belief tasks don’t accurately measure the social deficits of this group (in fact some adults with autism tend to pass first and second order false belief tasks, but this result cannot lead to the conclusion that they are necessarily normal in this domain; e.g., Ozonoff, Pennington & Rogers, 1991; Bowler, 1992).
The theoretical root of Baron-Cohen's studies (that we already can glimpse in his previous works and that was made clear in 1997 by Baron-Cohen, Wheelwright and Jolliffe) is the importance he gives to the ability to perceive a range of mental states, both cognitive and affective\(^{11}\), in interpreting and predicting the actions of others.

In 1997 he published two papers in which ToM were assessed using visual cues. The work by Baron-Cohen, Wheelwright and Jolliffe (1997) used photographs of an actress posing 10 basic emotions (happy, sad, angry, etc.) and 10 complex mental states (admire, interest, thoughtfulness, etc.), testing the role of face parts (eyes, mouth or whole face) in mindreading.

They found that for the basic emotions the whole face is more informative than single parts, while for the complex mental states, seeing the eyes alone produced the best performance; furthermore they pointed out that adults with autism or Asperger syndrome are impaired on the complex mental states recognition, in particular these subjects do not draw information from eyes, they have not a “language of the eyes”.

Probably, I suggest, it was the second work published in the same year by Baron-Cohen, Jolliffe, Mortimore and Robertson (1997) that influenced more deeply following ToM researches. In this paper they illustrated a very innovative ToM task for normal adults and adults with high-functioning autism and Asperger syndrome, called the “Reading the Mind in the Eyes”, or the “Eyes Test” for short, which items were first described by Baron-Cohen (1995), then revised in 2001a by Baron-Cohen,

---

\(^{11}\) Baron-Cohen and colleagues (e.g., Baron-Cohen, Wheelwright & Jolliffe, 1997; Baron-Cohen, Jolliffe, Mortimore & Robertson, 1997) gives the same meaning to the words “emotional” and “affective” mental states, so – as he does – I use in this paragraph the two terms interchangeably, even if I am aware (and I agree) that some authors, within the psychodynamic view (e.g., Fonagy et al., 2002), use differently these terms. In the next sections when I write about emotions I refer to discrete mental states, temporally and spatially defined, elicited from physics or psychological situation (Ekman, 1992; Ekman & Davidson, 1994), while the term “affective” denotes mental states that characterised all relationships, primarily caregiver-child interaction (Fonagy et al., 2002).
Wheelwright, Hill, Raste an Plumb (Italian normative data: Serafin & Surian, 2004). This final version of the test consists of 36 photographs (in the 1997 version there were 25 photographs) showing persons’ (male and female) eye region of the face. Participants are asked to make a forced choice, among which of four mental words (in the previous version of the test between 2 words) best describes what the person might be thinking or feeling. It consists also of a control tasks, the Gender Recognition Task (participants look at the same sets of eyes, but this time identifying the gender of person in each photograph), that does not imply mindreading; finally the test includes a Glossary, that subjects can read if they were unsure of the words meaning.

The Eyes Test was designed to be “a pure theory of mind test, at an advanced level” (Baron-Cohen et al., 1997, p. 816), that is it does not involve executive function (such as inhibition, planning) and central coherence components (there are not contextual information available). It is more than just an emotion perception test, because it includes terms describing also cognitive mental states; target mental states terms are, for example: upset, desire, fantasizing, regretful, doubtful, preoccupied, interested, serious, nervous, hostile, cautious, contemplative, suspicious.

They found that adults with autism/Asperger syndrome are impaired in this ToM test and that females perform significantly better than males (Baron-Cohen et al., 1997; 2001a). The Eyes Test can be considered a “real” measure of ToM, because it correlates to the Happè (1994) Strange Stories; furthermore it was demonstrated that adults with autism/Asperger syndrome performed as normal adults in the Gender Recognition Task and in an emotion task that assess basic emotion recognition from faces and they have the same level of general intelligence of normal populations: these data mean that the poor performance is not due to other deficits (e.g., in extracting social information from minimal cues, in basic emotion recognition, perceptual and intellectual deficits).
and that the mental states recognition from eyes can be an important aspect of social cognition (Baron-Cohen et al., 1997).

As the Authors suggested, this test is near to real social situations, obviously it simplifies the stimuli, that in everyday life are not static, but proceed rapidly, requiring the ability to grasp immediately (at an automatic and unconscious level) which intention, motives or other mental states they subtend; furthermore it involves the attribution of the relevant mental state (first stage of ToM attribution), but not of the content of it (that is why he/she has that mental state; it is the second stage of ToM attribution) (Baron-Cohen et al., 1997; 2001a).

The Italian version of the Eyes Test (named “Test degli Occhi”), by Serafin and Surian (2004), based on the reviewed version of the test (Baron-Cohen et al., 2001a), reports normative data for Italian adults. The authors, that administered the test to a normal population from 18 to 93 years old, pointed out two results, that were different from Baron-Cohen and colleagues finding: they did not find a gender differences in the test, but found that older persons have lower performance than younger adults, both in the Eyes Test and in the Gender Recognition Task.

Baron-Cohen, with Wheelwright, Spong, Scahill and Lawson (2001b) adapted the Eyes Test to create a child version of it (Italian translation: Liverta Sempio, Marchetti & Castelli, 2003) (Figure 2.1. shows an item of the Eyes Test – child version).

It comprises 28 photographs of the eye region of the face, chosen among those of the adult version of the test; children are asked to pick which of 4 mental words best describes what the person in the photo is thinking or feeling. Also in this version, the task includes the Gender Recognition Task, as a control for the non-mentalistic social attribution. It was administered to normal children and children with Asperger syndrome, age 6-10.
They found that children with Asperger syndrome are impaired in ToM and that, in normal population, older children performed better than younger children, revealing a main effect of age; they did not find gender differences.

Figure 2.1. An example of male eyes of the Eyes Test – child version (Baron-Cohen et al., 2001b); word choices were: hate, unkind, worried (correct), bored.

Auditory Theory of Mind tasks

Recently researchers’ attention have been also focused on another perceptive channel: the auditory one. Nearly at the same time, two research groups created, independently between them, two advanced ToM instruments, for adults with autism, based on vocal cues.

In 2001, Kleinman, Marciano and Ault created an auditory ToM task, the “Mental State Voices Task (MSVT)”, consisting of a sentence without inherent emotional content (“The quick brown fox jumped over the lazy dog”), recorded using different intonations, in order to represent six basic emotions (happy, sad, angry, afraid, surprised, disgusted) and six complex mental states (arrogant, guilty, calm, anxious, bored, interested). Participants choose between to adjective which best describe what the actor was feeling or thinking.
They used, like in the *Eyes Test*, a gender control task using the same voices of the MSVT. Their aim was to compare, in adults with autism, performance in an advanced visual mental state attribution (the *Eyes Test* by Baron-Cohen, Wheelwright & Jolliffe, 1997) with that in the MSVT, considered as a parallel auditory task. Their work extended Baron-Cohen, Wheelwright and Jolliffe’s (1997) work by showing that the deficit in individuals with high-functioning autism is not limited to the “language of eyes”, but it includes also the identification of complex mental states from voice. They found that participants were significantly more accurate in assessing mental states based on the voices as contrasted with the eyes task, concluding that “for people to assess emotions, disembodied verbal information appears superior to information gleaned from the eyes” (Kleinman, Marciano & Ault, 2001, pp. 33-34); they failed to find a gender differences in both perceptive tasks.

In the same year, Rutherford, Baron-Cohen and Wheelwright (2001) presented a research with the same aim of Kleinman, Marciano and Ault (2001). In order to expand Baron-Cohen and colleagues’ (2001a) finding about the adults with autism/Asperger syndrome from the visual domain into the auditory one, they create a new task: the “Reading the Mind in the Voice Test”, shortly the “Voice Test”.

It consists of 40 segments of speech, selected from 50 samples of dialogue recorded from dramatic audio books (e.g., “Collie said you were up here”; “Please! We must go”). Subjects choose between two adjectives to best describe the mental state of the speaker (e.g., referring to the two examples cited above, “friendly” as target term and “grateful” as foil; “worried” as target and “insulted” as foil).

Each segment last for approximately 2 seconds with a 3-seconds pause between sentences, during with participants marked their choice.
The *Voice Test* investigates mentalistic attributions, for example target mental states included in the test are: friendly, suspicious, embarrassed, sincere, reassuring, sarcastic, reflective, hopeful, annoyed, resigned, disappointed, worried. The control task consists of the same voices of the test and participants judge the speaker’s age (under or over 42).

They found that people with high-functioning autism/Asperger syndrome have a specific deficit making social inferences in auditory domain, too. The *Voice Test*, like the *Eyes Test* (Baron-Cohen et al., 1997; 2001a), do not correlate with IQ: the complex mental states recognition from perceptive cue is independent of intellectual ability.

The Voice Test was recently revised12 (Golan, Baron-Cohen, Hill & Rutherford, 2007a), because authors found some limitations: it had only two possible answers per question (while the task’s sensitivity could be improved by adding foils to each item) and some items could not discriminated autism spectrum condition group from the normal one. So they removed some items and added foils (selected from the emotion taxonomy created by Baron-Cohen, Golan, Wheelwright & Hill, 2004, consisting in 412 mental states, each in one of six developmental levels). The final version of the test consists of 25 item, presented in a random order, using a computer software. It has good reliability and validity, it is harder and more sensitive than the first version (Rutherford et al., 2002) in distinguishing adults with autism/Asperger syndrome from normal adults.

They also found that the verbal IQ has a significant effect on the task scores and they explain this result referring to the difficulty to choose among several foils, that probably require the involvement of verbal ability to distinguish several potentially correct mental states’ verbal labels in order to find the most suitable one.

---

12 The revision of the Voice Test (Golan et al., 2007) was published after the present research was done, so the Italian child version of the Voice Test was created independently from the recent English revision.
They did not find differences between male and female in normal adult population and found a weak, even if significant, correlation between the *Voice Test* and the *Eyes Test*, in fact the two tests are related to different perceptual modalities and require the recognition of quite different emotions and mental states (Golan et al., 2007).

**Visual and auditory Theory of Mind tasks**

It could be useful, as already noted Kleinman, Marciano and Ault (2001), to determine how the joint contribution of facial and verbal information together affects accurate processing of mental states.

Now I briefly describe two new tests, worked out in 2007, that use both visual and auditory channel, getting closer to real social situations, in which rarely faces and voices are separated. These tests were administered to individuals with autism spectrum conditions, in order to analyze deeper their ToM deficit.

Golan, Baron-Cohen and Hill (2006a) constructed a battery of visual and complex emotion recognition tasks, called the “*Cambridge Mindreading Face-Voice Battery*” (CAM), that evaluates, separately, the recognition in the face (Figure 2.2.) and in the voice of 20 emotional concepts, each expressed by 5 items (2 or 3 faces and 2 or 3 voices) from the taxonomy of emotion (Baron-Cohen, Golan, Wheelwright & Hill, 2004). After watching a clip or listening a voice, the participant chooses the word, among 4 adjectives, which better describe how the person is feeling. They showed that the vocal and the visual scales strongly correlated and that males with autism spectrum conditions had more difficult to recognize emotions from faces than from voices (while this difference was not found in females and in the control group) This battery has the merit to test the same emotions, using different perceptive channels.
Golan, Baron-Cohen, Hill & Golan (2006b) created a complex emotion and mental states recognition task, called the “Reading the Mind in Films” task (RMF).

It consists of 22 short social scenes, taken from films, including visual input (facial expression, body language, action), auditory input (prosody, verbal content) and context (e.g., a scene taken from “The Turn of the Screw” (1999): a young woman complimenting an older woman on the way she educated the children; the older woman thanks her calmly, then runs towards her with tears in her eyes saying: “Oh miss, I’m so glad you’re here”; target term is “overcome”, foils are “sociable”, “admiring”, “liked”). It requires the integration of multimodal information. Participants were asked to identify the protagonist and to label his/her emotion or mental state at the end of the scene, choosing it among 4 words. They found that individuals with autism spectrum conditions performed lower than controls and that the task correlated with verbal IQ, suggesting that participants used verbal content to pick up the protagonists’ mental states.

The authors used also the various perceptive tasks to teach individual with autism spectrum condition to improve their mental states recognition (e.g., Golan & Baron-Cohen, 2006).

**Figure 2.2.** Example of item of the emotion recognition in the face task (Golan, Baron-Cohen & Hill, 2006a); word choices were: restless, sadistic, Cherishing, flattering.
Some final suggestions

In conclusion of this historical *excursus* within ToM studies, I briefly underline some aspects. From the methodological point of view, it is evident the growing complexity, during the years, of instrument for perceptual ToM assessment: from pictures to real faces, from static stimuli to dynamic stimuli, from the separation to the integration of multimodal information. Tasks become therefore more and more ecologic, near to real daily life.

A limit of the works I presented regards the population they were created for; in fact all the studies were aimed to better understand the autistic population and not the normal one (even if normal population were tested as control group), so their difficulty was calibrated on a pathological population, that studies revealed to have difficult to understand mental states (e.g., in classical false belief tests, in narrative advanced tasks and in tasks about ToM precursors) (e.g., Happè, 1994; Baron-Cohen, 1995; Surian, 2002; Frith, 2003; Tager-Flusberg, 2007). So items that constitutes these tasks are easier for normal population than for clinical one.

Furthermore they were administered usually on adults, only few tasks were created specifically for the children population, so they cannot analyze the developmental course of ToM after the false belief comprehension.

From the theoretical point of view, I suggest that the creation of these tasks testifies the idea that ToM is a continuous and multi-faceted ability and persons can have different levels in various aspects of the mindreading ability. The so called normal population is able to use, automatically and contemporaneously, different channels to understand mental states and probably they can provide with some high ToM abilities for a possible low level in other specific ToM abilities, so they appear competent in social relationships.
In the developmental context, this ToM view lead researchers to study which and when specific mentalization abilities develop and how (which communicative channel) children grasp mental states during daily life.

The *Voice Test* for the Italian normal children population was created with the aim to analyze deeper ToM (understood as a multi-componential ability, that encompasses a wide range of epistemic and emotional mental states) development in school age children. In fact Italian ToM researchers can count on the translation (but not Italian validation and standardization) of some advanced ToM tasks, regarding the cognitive and the visual channel, but not the auditory one.

The *Voice Test* child version I present in the next section is a new test, developed in Italian (it is not a simply translation of an English task) (phase 1), validated (phase 2) and standardized (phase 3) in a normal children population, from 7 to 11 years old.

### 2.2. Phase 1: Construction of the Voice Test

The *Voice Test* was developed following three main steps. First: the creation of the items and the preparation of the auditory stimuli. Second: the reduction of items on the basis of the correspondence between auditory stimuli and mental state terms and of the semantic neutrality; then it was run the analysis of three different administration procedures. Third: the analysis of items difficulty and discrimination and the internal consistency of the test and the consequent further items reduction, in order to produce the *Voice Test* final version.
2.2.1. Preparation of the stimuli

On the variety of ways to approach construction of the Voice Test items to understand mental states in the voices, the one adopted here was to use adult actors to produce stimulus items. This test procedure, used successfully in previous works within the field of study of the emotion paralinguistic cues (see, for example, Baum & Nowicki, 1998; Rothman & Nowicki, 2004), has the advantage that the production of items was guided by researchers’ direction. In fact, initially 48 items were created; each item was constituted of a segment of dialogue, recorded by professional actors and actress and of four terms regarding complex epistemic or emotional mental states, written on the answer sheet. Children were asked to pick which of the four words best describes what the speaker’s mental state (only one term, among the four words, correctly represents what the speaker is thinking or feeling). The four options were developed on the basis of the other ToM perceptive tasks: the Reading the Mind in the Voice test (Rutherford, Baron-Cohen & Wheelwright, 2002) for adult and the Eyes Test – Child version (Baron-Cohen et al., 2001b) translated in Italian (Liverta Sempio, Marchetti & Castelli, 2003).

Each set of complex mental terms, referred to epistemic and emotional mental states, was built considering the school age children’s social experience. In particular, the terms that in the Voice Test by Rutherford, Baron-Cohen and Wheelwright (2002) were typical of adult life (e.g., flirtatious) were not included in the test. Besides mental terms taken from the Voice Test and the Eyes Test – Child version, it were included others mental states typical of children experiences (e.g., pleading). So there were 48 target mental states, each associated to three foils.

The segments of dialogue of each item were identical to those, considered suitable for children, used in the Reading the Mind in the Voice task (Rutherford, Baron-Cohen &
Wheelwright, 2002) or were specially created for the present task, following the structure of Rutherford, Baron-Cohen and Wheelwright’s (2002) sentences. All the segments of dialogue were semantically neutral, in fact they can be spoken in a number of different ways to communicate a specific mental states at various level of intensity. So each item refers to an epistemic or emotional mental state depending on the way it is pronounced by the actor/actress (non verbal cues), independently from the semantic content of the phrase (for example, the segment “I know you met Tina” is pronounced in a particular way – using a specific tone of voice, rhythm… – associated to the target mental term “thinking to something”).

A Glossary with explanations and examples for all the mental words were built, that participants can look it up if they need an explanation of mental terms. The Glossary was built following the glossaries created for the other ToM tasks based on perceptive cues, from which the mental states were taken (i.e., the Reading the mind in the Voice and the Eyes Test – Child version) (see Appendix).

The translation process of the mental state terms, the spoken phrases from the adult version of the Voice Test (Rutherford, Baron-Cohen & Wheelwright, 2002) and the Glossary followed the standard guidelines (Hambleton, 1994); two Italian translators, perfect English speakers and expert in the field of ToM, translated the items, then an English native speaker back-translated them and compared her version to the original version.

Stimuli were recorded using the voices of two females (one under 25 and one over 60 years) and two male (one under 25 and one over 60 years) Italian adults, with current theatrical acting experience. Sentences were recorded on an audio cassette. Each speech segment lasted for approximately 2 second, with a 3-second pause between speech segments.
This test requires to choose the correct answer, among four options, while the English
*Voice Test* (Rutherford, Baron-Cohen & Wheelwright, 2002) involved a forced choice
between only two response options (but the chance performance on each trial was
high: $p = 0.5$). The task’s sensitivity (as noted also recently Golan, Baron-Cohen, Hill &
Rutherford, 2007) could have been improved by adding foils, so the Italian *Voice Test*
had two more foils for each item (similarly to the revised version of the *Eyes Test* by
Baron-Cohen et al., 2001a, and the child version of the *Eyes Test* by Baron-Cohen et al.,
2001b). Moreover, taking (target and foils) mental states from existing validated ToM
tests satisfied the content validity of this new test: the items were representative of the
mental states the test was originally designed to measure (Boncori, 2006).

For control trials, the answer sheet had two options to represent the person’s gender
(“male” and “female”) and two option to represent the person’s age (“under 25” and
“over 60”). The original English version of the test asked only the person’s age, the
gender recognition task was here added because children, having less life experience
than adults, can find difficult to recognize persons’ age (for example, to recognize the
possible change in voice features during the course of life) and in order to make similar
the child version of both the *Voice Test* and the *Eyes Test*. The entire procedure I
describe in this paragraph, to conceive the *Voice Test* stimuli, was developed by O.
Liverta Sempio, and A. Marchetti of the Theory of Mind Research Unit, Catholic
University of the Sacred Heart, Milan.

### 2.2.2. Item reduction and administration procedure

Four independent adult judges were asked to judge the speaker’s mental state, among
the four options, as they listened to the audiotape; they took the test in real time, with
a 3-second pause between segments. Seven items on which the judges were not unanimous were kept, leaving 41 segments of speech. The same procedure was followed for the control task: again, four independent judges were given the questionnaire and asked to judge speaker’s age and gender, between the two options, as they listened to the audiotape; all the judges answered correctly to this task (Liverta Sempio, Marchetti & Fabio, 2005).

To verify if the 41 items were neutral in their meaning, 22 children (13 male and 8 female, coming from a middle-class background, which parents gave their consent to participate to this research task) without learning problem, attending the third class in a primary school near Milan, were asked to judge if the sentences expressed an intense negative or positive meaning, or a neutral, or a very little negative and positive meaning. This procedure selected sentences which meaning did not influence children’s mental state attribution to the speaker, in order to make the Voice Test a task that assesses the ability to understanding mental states only from nonverbal vocal cues (paralinguistic signals).

Each child read by him/herself the sentences written on the answer sheet; 4 control sentences (with a very negative and positive meaning) were added to the 41 items and they were correctly judged from all the children. Four target sentences did not receive a judgment of neutrality (including the judgment of neutral, very little negative and very little positive meaning) from the 75% of children and were removed from the test, leaving 37 items.

Before the Voice Test creation proceeded any further, it was verified if the test could be administered, without altering children’s performance, also in group. This step was important because a group setting reduce the time for administration (more data collection in less time) and allows to use easily the test in different contexts (for
example not only experimental one, but also school context). To this aim children were tested in different settings.

Participants

Sixty-six children (33 males and 33 females), born in 1998 (mean age: 88.20 months; standard deviation: 3.30 months), were recruited from four primary schools located near Milan; they all attended the second class and came from a middle-class background. Two subjects (a male and a female) were removed because they had learning problems.

All children’s parents gave their written consent to allow children’s participation in this study.

Materials and Procedures

Children were administered the 37 items of the Voice Test (preceded by the trial item), in a quiet room at their school; they were asked to pick which of 4 words best describes what the speaker is thinking or feeling. Immediately before listening each items, participants were shown and read the 4 options on the answer sheet and asked whether they did not understand any words or they were unsure of any words meaning. If they needed an explanation of a word, the experimenter read aloud some synonymous and an example from the Glossary. The experimenter paused the audiotape between speech segments to give children the time they needed to choose the mental state term.

After the administration of the test, they listened again to the same 37 items (preceded by the trial item) and they were given the control task (recognition of the speakers’ gender and age): they marked if the speaker was a male or a female and if he or she was
young (under 25 years) or old (over 60 year). Children were allowed to ask to adjust the volume as necessary before and during the test.

The same procedure (see Appendix for detailed instruction) was also used in the following administration.

Participants were tested in three different setting:

- 20 children (10 males and 10 females) were tested individually;
- 24 children (12 males and 12 females) were tested in little groups of three children (four groups made up of two males and a female and four groups of two females and a male);
- 20 children (10 males and 10 females) were tested in group of 10 children (5 males and 5 females in each group).

**Results**

An analysis of variance was conducted to compare children’s performance in the three administration settings. The one-way ANOVA with the condition of administration as the independent variable and children performance in the Voice Test as dependent variable did not show significant differences between means for different setting: F (2, 61) = 1,123, p = 0,33.

All the children passed the gender recognition task. To analyze if children were able to recognize the speakers’ age, a frequency analysis was run for each item. Thirty-one items received by the 40-60% of the participants the correct answer, only six items were correctly recognized by most of the children (>75%).

An ANOVA was run to analyze if the administration setting influenced the recognition of the speakers’ age, but it did not reveal significant differences: F (2,61) = 0,565, p = 0,571.
Discussion

The performance in the Voice Test in the three different administration setting (individual, little group of three children, group of ten children) was not different; the following administration therefore could be done in group of 10 children.

About half of the children failed in the age recognition control task and this result was not due either to the administration context or to a lack of attention or auditive problems, because they correctly identified the speakers’ gender. So this task was removed from the test because of its difficulty for children and only the gender recognition task remained as a control task, like in the Eyes Test.

2.2.3. Further item reduction: the final version of the Voice Test

The first step was to sample the 37 items onto a computer and digitally clean tape recording noise. The task was therefore presented, from now on, to the participants using a CD player. Moreover the answer sheet for control trials was changed in order to represent only the two gender options (“M” for male and “F” for female).

In order to create a valid test, it was necessary to identify the items that were too much simple or difficult; to analyze internal consistency of the Voice Test; and to verify if children’s accuracy in identifying the mental state in voices increased with age.

Participants

170 children (80 male and 90 female; mean age: 100,09 months; s.d.: 16,59 months), aged between 73 and 131 months, were recruited from five primary school located near Milan. They received their parent’s written consent to participate in the study. They were all Italian native speaker and came from a middle-class background.
They were subdivided in 5 groups, corresponding to the class they were attended:

- first class: 29 children (14 males and 15 females) from 73 to 84 months; mean age: 77.34 months (s.d.: 2.98);

- second class: 32 children (16 males and 16 females) from 83 to 93 months; mean age: 87.95 months (s.d.: 3.41);

- third class: 38 children (13 males and 25 females) from 97 to 108 months; mean age: 101.58 months (s.d.: 3.40);

- fourth class: 38 children (18 males and 20 females) from 108 to 120 months; mean age: 113.61 months (s.d.: 3.78);

- fifth class: 33 children (19 males and 14 females) from 120 to 131 months; mean age: 125.85 (s.d.: 3.75).

*Materials and Procedures*

Children were tested in groups consisting of a maximum of 10 children (children attending the first class were tested in group of three, in order to better control their understanding of the test procedure), in a quiet room in their school.

They listened to the voices, recorded on the CD, and were asked to make a judgment on each speech segment, choosing a mental state term among 4 words. The *Voice Test* answer sheet and the instruction were the same written in the previous section (see also Appendix); the answer sheet of the control trials required only to mark if the speakers were male or female, so children were asked to make gender judgments for the same segments on which they had made mentalistic attribution before. While in the Voice Test, the experimenter paused the CD-tape between segments (without repletion of the items), in the gender recognition task children took the test in real time, in fact,
according to the experimenter observation during previous administration, 3 seconds was sufficient time for every subject. The Voice Test and the control task took about 20 minutes.

Results

All the children answered correctly to the gender control task.

The statistics typically used in the item analysis of tests are those referred to item facility/difficulty, internal consistency of the test and item discrimination.

The index of difficulty was calculated for all the individual items using the Guilford’s (1956) difficulty index: \( \text{Pc} = \frac{(np - 1)}{(n - 1)} \). \( \text{Pc} \) is the corrected proportion (it corrects the formula of \( p \), that is the proportion of children who answered the item correctly) and it can take a score included between 0 (very facile) and 1 (very difficult). In this work it was chosen an intermediate level of difficulty: \( \text{Pc} = 0,50 (\pm 0,20, \text{that is} 0,30 < \text{Pc} < 0,70) \).

The reason for measuring item difficulty is to choose items of suitable difficulty level which help in assessing as accurately as possible each individuals level of knowledge. Thus an item with an index of difficulty higher than 0.70 and lower than 0,30 is deemed to be a poor discriminator.

Guilford’s indexes for each item are presented in Table 2.1. (see also Chart 2.1.). The items did not satisfy the intermediate level of difficulty were 15 and they were removed from the Voice Test; the test was therefore constituted of 22 items.

---

13 In the Guilford’s formula “n” represents the number of options, in the case of Voice Test it was 4; “p” is the proportion of correct answers (“number of correct responses / total number of response”).
Table 2.1. Guilford’s difficulty index (Pc) for each item (selected items are highlighted)

<table>
<thead>
<tr>
<th>Item</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>8</th>
<th>9</th>
<th>10</th>
<th>11</th>
<th>12</th>
<th>13</th>
</tr>
</thead>
<tbody>
<tr>
<td>Pc</td>
<td>0.2</td>
<td>0.5</td>
<td>0.1</td>
<td>0.5</td>
<td>0.1</td>
<td>0.4</td>
<td>0.7</td>
<td>0.5</td>
<td>0.3</td>
<td>0.2</td>
<td>0.6</td>
<td>0.7</td>
<td>0.7</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Item</th>
<th>14</th>
<th>15</th>
<th>16</th>
<th>17</th>
<th>18</th>
<th>19</th>
<th>20</th>
<th>21</th>
<th>22</th>
<th>23</th>
<th>24</th>
<th>25</th>
<th>26</th>
</tr>
</thead>
<tbody>
<tr>
<td>Pc</td>
<td>0.0</td>
<td>0.5</td>
<td>0.5</td>
<td>0.3</td>
<td>0.7</td>
<td>0.6</td>
<td>0.1</td>
<td>0.7</td>
<td>0.4</td>
<td>0.1</td>
<td>0.4</td>
<td>0.1</td>
<td>0.0</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Item</th>
<th>27</th>
<th>28</th>
<th>29</th>
<th>30</th>
<th>31</th>
<th>32</th>
<th>33</th>
<th>34</th>
<th>35</th>
<th>36</th>
<th>37</th>
</tr>
</thead>
<tbody>
<tr>
<td>Pc</td>
<td>0.4</td>
<td>0.7</td>
<td>0.1</td>
<td>0.0</td>
<td>0.4</td>
<td>0.4</td>
<td>0.6</td>
<td>0.0</td>
<td>0.2</td>
<td>0.1</td>
<td>0.0</td>
</tr>
</tbody>
</table>

Chart 2.1. Guilford’s difficulty indexes
The determination of the internal consistency for the 22 items of the Voice Test was accomplished through the Cronbach’s coefficient \( \alpha \). Cronbach’s alpha was 0.60; in order to improve alpha coefficient two item were removed, because they were poor contributors to the overall measure. Internal consistency for the 20 items was 0.69.

Finally, it was analyzed if Voice Test differentiates correctly among different ages. The total score was distributed normally (Chart 2.2.), with a mean of 13.18 (s.d: 3.27), a median of 14.00 and a mode of 12.00. The mean of the total error score was 6.82 (s.d.: 3.27). The correlation between chronological age and total number of errors on the test for all participants was significant: \( r = -0.541, p < 0.001 \).

Table 2.2. shows means and standard deviations of the total error scores and total scores of participants, subdivided in five groups of age; Chart 2.3. represents the Voice Test total score increase with age.

An ANOVA was run, with age (the five class) as independent variable and the Voice Test total score as dependent variable: \( F (1, 4) = 19.99, p < 0.001, \eta^2 = 0.33 \).

<table>
<thead>
<tr>
<th>Mean age (months)</th>
<th>Total error mean (s.d.)</th>
<th>Total score mean (s.d.)</th>
</tr>
</thead>
<tbody>
<tr>
<td>77.34 (1\textsuperscript{st} class)</td>
<td>8.97 (3.05)</td>
<td>11.03 (3.05)</td>
</tr>
<tr>
<td>87.95 (2\textsuperscript{nd} class)</td>
<td>8.81 (3.22)</td>
<td>11.19 (3.23)</td>
</tr>
<tr>
<td>101.58 (3\textsuperscript{rd} class)</td>
<td>7.45 (2.46)</td>
<td>12.55 (2.46)</td>
</tr>
<tr>
<td>113.61 (4\textsuperscript{th} class)</td>
<td>4.84 (2.16)</td>
<td>15.16 (2.16)</td>
</tr>
<tr>
<td>125.85 (5\textsuperscript{th} class)</td>
<td>4.58 (2.72)</td>
<td>15.42 (2.73)</td>
</tr>
</tbody>
</table>
Chart 2.2. The normal distribution of the Voice Test total score

![Chart 2.2](image-url)

- **Mean** = 13.1765
- **Std. Dev.** = 3.27145
- **N** = 170

Chart 2.3. The Voice Test score increments with age

![Chart 2.3](image-url)

- **Voice Test total score**
- **Age (class)**
Discussion

Briefly, preliminary items analyses were performed in order to obtain the initial validation of the *Voice Test*. The final version of the test consisted of 20 item (see Appendix); 17 items were removed after the analyses of item difficulty (Guilford index) and of the internal consistency. The test showed an acceptable estimate of internal consistency (Cronbach’s *alpha* was 0.69) (Cohen, 1960).

It was also verified that the test scores were normally distributed and that children’s performance increased with age, from 6.4 to 10.5 years. The results indicated that the psychometric qualities of the Voice Test final version appear to be sound.

2.2.4 Conclusion

The current study describes the three steps of the development of the *Voice Test*, an advanced Theory of Mind test based on auditory cues.

From the 48 initial items, consisted of speech segments representing specific complex mental states acted by two Italian actors and two Italian actress, were excluded 28 items, in order to create a quick, easy to use and valid test.

A series of analyses were run in order to verify: the correspondence between auditory stimuli and mental state terms; the speech segments semantic neutrality; the suitability of the group administration procedure and of the control tasks; the items difficulty; the items discrimination among five groups of age (within children attending the primary school, from the first to the fifth class); the normal distribution of the *Voice Test* total score; and the internal consistency of the test.

The final version of the *Voice Test* consists of 20 items (preceded by the trial item), of which 10 acted by two female voices and 10 by two male voices, representing both
epistemic and emotional complex mental states, followed by a control task implying a social but non-mentalistic ability (gender recognition task). It also includes a Glossary that helps children to better understand mental state terms meanings.

Children were asked, for each item they listened, to pick which of 4 words best describes what the speaker is thinking or feeling. The Voice Test, in its final composition (see Appendix), takes about 20 minutes (30 minutes with younger children and 15 minutes with older children).

The test can be really considered an advanced ToM test for children because it assesses the understanding of a wide range of mental states from vocal cues; for example it goes beyond the concepts of desire and belief including the comprehension of more developed mental processes such as “thoughtful” and “confused”, and it assesses the recognition ability of not-basic emotion, such as “nervous” or “worried”.

The Voice Test is therefore an ecological test that differs both from other ToM tests (e.g., Happe, 1994; Baron-Cohen et al., 2001a, 2001b; Rutherford, Baron-Cohen & Wheelwright, 2002) – because of the stimuli it is constituted (voices) and of the population it refers (normal children) – and from other auditory tests regarding emotional competence (e.g., Rothman & Nowicki, 2004) – because of the wide range of mental states it assesses.

In conclusion, the preliminary analyses demonstrated the acceptability of the administration group (maximum of 10 children) procedure, of the content validity and of the psychometric properties of the Voice Test.

The next paragraphs present the validation and standardization of the test in a large children population sample.
2.3. Phase 2: Validation of the Voice Test

Validity refers the extent to which a test measures what it is supposed to measure (“Is the test measuring what you think it is measuring?). It is a judgment made on the basis of experience and empirical indicators (Boncori, 2006). In this work it were analyzed the content validity and the construct validity of the Voice Test.

Content validity is a non-statistical type of validity that involves the examination of the test content to determine whether it is a representative sample of the specific domain to be measured. This type of validity id ordinarily to be established deductively (Cronbach & Meehl, 1955), I describe in the paragraph 2.2.1. the content validity of the Voice Test.

Construct validity refers to the degree to which inferences can legitimately be made from the operationalizations done in the study to the theoretical construct on which the operationalizations were based (Cronbach & Meehl, 1955; Pedrabissi & Santinello, 1997).

In order to verify this type of validity, it were analyzed the convergent validity of the Voice Test, that is the degree to which the Voice Test is correlated with other ToM measures (e.g., the Voice Test is theoretically predicted to correlate with another ToM perceptive test).

The aim of the present study was to validate the new advanced ToM instrument based on auditory cues, the Voice Test, on an Italian sample constituted of normal children attending the primary school.
2.3.1. Method

Participants

The sample consisted of 220 children (111 males and 109 females), aged between 78 and 134 months (M = 105.17 months; s.d. = 16.73), attending the primary school in three towns near Milan. Children came from a middle-class background. They had not psychological or neurological proclaimed pathologies and learning difficulties. A few participants (less than 3%) were not Italian native speakers, but all were competent in speaking and understanding Italian. The participants’ parents gave their written consent to allow children’s participation in this study.

Participants were subdivided in five groups of age:

- 54 children (25 males and 29 females): 7 years old (78-89 months);
- 41 children (23 males and 18 females): 8 years old (90-101 months);
- 48 children (25 males and 23 females): 9 years old (102-113 months);
- 46 children (27 males and 19 females): 10 years old (114-125 months);
- 31 children (11 males and 20 females): 11 years old (126-137 months).

Measures

The Voice Test is an advanced ToM task consisted of 20 auditory stimuli (the test creation is described in the paragraph 2.2.; see also Appendix). It assesses ability to read complex epistemic and emotional mental states from voices. Children were asked to judge what the speaker is thinking or feeling (his/her mental state), from his/her voice alone, choosing a word among four mental state terms; then they judge the same speakers’ gender (male or female).
Children were tested on three ToM tasks: two classical first-order (*The Deceptive Box*) and second-order (*Look Prediction*) false belief tasks and an advanced measure of ToM ability, based on perceptive cues, the *Eyes Test* – child version.

The *Deceptive Box Task* (Perner, Leekam & Wimmer, 1987; Italian adaptation: Liverta Sempio & Machetti, 2001c) assesses the first-order false belief. Child was shown a closed box of pastels (it was a well-known brand of pastels) and asked what they thought was inside. The box was then opened to show the child that it contained not pastels, but a little doll. The doll was replaced and the child was then asked a memory control question: “What is inside here?”. Then the child was asked to predict what a friend of his/her, who did not look into the box because he/she was not in the room, would think was in the box. After this question, regarding the false-belief of another person, the experimenter asked a question about child’s false belief: “Before I opened this box, what did you think was inside?”.

The *Look Prediction* is a second-order false belief task (Sullivan, Zaitchik, & Tager-Flusberg, 1994; Italian adaptation: Antonietti, Liverta Sempio, Marchetti & Astington, 1999). Children were told a story, with four pictures helping them to follow the plot, about John and his sister Mary: John and Mary are playing in John’s bedroom with a new pack of cards, when Mary tells her brother she is going to put cards into the wardrobe, because she will go to the kitchen to lay the table; when Mary goes out the bedroom, John decides to play a trick to her: he takes the cards from the wardrobe and puts them under his bed. Children were asked two control questions: “Did Mary see John hiding the pack of cards under his bed?” (memory question) and “Where does Mary think is the pack of cards?” (reality question). After passing these questions, they were asked a first-order false belief question (“Does John think that Mary saw him hiding the pack of cards under his bed?”) and a second-order false belief question:
“Where does John think that Mary looks for the pack of cards when she comes back into his room? And why does she look for there?”

The answers in these two false belief tasks were scored dichotomously as pass-fail; for every correct answer children got one point, for a maximum of four points in each task. The Eyes Test – child version (Baron-Cohen, Wheelwright, Spong, Scailhill & Leason, 2001b; Italian version: Livert Sempio, Marchetti & Castelli, 2003) assesses the ability to recognize and attribute mental states from visual cues. It consists of 28 photographs of the eye region of the face; children were asked to pick which of four words best describes the person in the photo is thinking or feeling; only a word is correct. Children were also administered a control task, in which the same 28 photos were shown and they were asked to judge the person’s gender. For each correct answer, children got one point. Information on reliability and validity is not available for the test, but Baron Cohen and colleagues (2001b) found that the total score improve across the age-groups (6-8; 8-10; and 10-12 years) and ANOVA testing revealed that the Eyes Test was highly significant in distinguishing between differentiated performance of these groups. The adult version of the test (both in English and in Italian; Baron-Cohen et al., 2001a; Serafin & Surian, 2004) is validated.

Finally, as verbal measure, it was administered a test of verbal mental age: the Peabody Picture Vocabulary Test – Revised (PPVT-R) (Dunn & Dunn, 1981; Italian validation: Stella, Pizzoli & Tressoldi, 2000); it is a measure of receptive vocabulary for standard Italian, it is a norm-referenced power test containing 5 training items followed by 175 test items arranged in order of increasing difficulty.
Procedure

Children were tested over two session spaced on two weeks apart, in a quiet room in their school. At session 1, children were given in group of ten children the *Voice Test* and individually the PPVT-R and the first-order *False Belief* task. At session 2, children were tested individually on the second-order *False Belief* task and the *Eyes Test* – child version.

2.3.2. Results

Descriptive statistics

Table 2.3. presents means and standard deviations for all variables. The first-order *False Belief* task was not passed only by eight children; the first-order question in the second-order *False Belief* task was passed by 213 children, while the second-order question by 178 children; these tasks had not a normal distribution. The first-order false belief task did not included in following analyses, while it was used the non parametric statistics when the second-order *False Belief* task was considered in the analysis. All the other tasks were normally distributed.

All the children participating in this study passed the control questions of the *False Belief* tasks and the *Voice Test*; 115 children recognized correctly the person’s gender in the control task of the Eyes Test, however all the children recognized correctly at least the gender of 20 persons out of 28 (mean: 26,61; s.d.: 1,79).
Table 2.3. Means and standard deviations

<table>
<thead>
<tr>
<th>Test</th>
<th>M</th>
<th>s.d.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Voice Test (0-20)</td>
<td>13.10</td>
<td>3.51</td>
</tr>
<tr>
<td>Eyes Test (0-28)</td>
<td>15.85</td>
<td>4.14</td>
</tr>
<tr>
<td>First-order False Belief (0-4)</td>
<td>3.95</td>
<td>0.25</td>
</tr>
<tr>
<td>Second-order False Belief (0-4)</td>
<td>3.79</td>
<td>0.46</td>
</tr>
<tr>
<td>PPVT – R</td>
<td>101.94</td>
<td>15.00</td>
</tr>
</tbody>
</table>

**Influence of gender and age**

The independent-sample *t* test did not show gender differences in the *Voice Test*, in the *Eyes Test* and in PPVT – R.

The nonparametric test, Mann-Whitney *U* Test, was run to investigate gender differences in the second-order *False Belief* task and showed a significantly difference between male and female (female performance better than male): *U* (218) = 4937.500, p = 0.001, two-tailed.

Both the *Voice Test* and the *Eyes Test* score increased with age: *F* (4, 215) = 34.075, *p* < 0.001 and *F* (4, 215) = 15.301, *p* < 0.001. The LSD post-hoc test revealed that between 9 and 10 years old and between 10 and 11 years old, the performance in the *Voice Test* did not increase significantly (Chart 2.4.). The PPVT – R increased with age [*F* (4, 215) = 8.694, *p* < 0.001], but the LSD post-hoc revealed that differences were only between 6-7 and 7-8 years; there were not differences among 9, 10 and 11 years old. The Mann-Whitney *U* Test was run to analyze if the performance in the second-order *False Belief* increased with age: differences among age were significant (*p* < 0.01), except among 9, 10 and 11 years old.
Chart 2.4. Age influence on the Voice Test performance.

**Bivariate correlations**

Person’s correlation coefficient $r$ was used to analyze correlations between normally distributed variables, while its nonparametric version, Spearman’s $\rho$, was used when the second-order False Belief (2nd FB) score was considered. Table 2.4. presents the bivariate correlations between variables. In detail, the *Voice Test* correlated positively with the *Eyes Test* (Chart 2.5.), the second-order *False Belief* task (Chart 2.6.) and the receptive language ability (PPVT-R) (Chart 2.7.).

Controlling for verbal ability the correlation between the *Voice Test* and the *Eyes Test* remained significant: $r = 0.404, p < 0.001$. Also controlling for age it was significant: $r = 0.363, p < 0.001$. Finally, the partial correlation controlling both for age and PPVT – R was significant: $r = 0.269, p = 0.002$.

Controlling for age the correlation between the *Voice Test* and the PPVT – R remained significant: $r = 0.329, p < 0.001$. 
**Chart 2.5.** Correlation between the Voice Test and the Eyes Test

![Chart](chart.png)

**Table 2.4.** Correlations (the Voice Test correlations are highlighted)

<table>
<thead>
<tr>
<th></th>
<th>Eyes Test</th>
<th>2nd FB</th>
<th>PPVT – R</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Voice Test</strong></td>
<td>0.532</td>
<td>0.357</td>
<td>0.408</td>
</tr>
<tr>
<td></td>
<td>*p &lt; 0.001</td>
<td>*p &lt; 0.001</td>
<td>*p &lt; 0.001</td>
</tr>
<tr>
<td><strong>Eyes Test</strong></td>
<td></td>
<td>0.352</td>
<td>0.422</td>
</tr>
<tr>
<td></td>
<td></td>
<td>*p &lt; 0.001</td>
<td>*p &lt; 0.001</td>
</tr>
<tr>
<td><strong>2nd FB</strong></td>
<td>1</td>
<td></td>
<td>0.144</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>*p&lt;0.05</td>
</tr>
</tbody>
</table>
Chart 2.6. Correlation between the 2nd order False Belief task and the Voice Test

Chart 2.7. Correlation between the Voice Test and the PPVT – R.
2.3.3. Discussion and conclusion

The present study investigated the construct validity of the *Voice Test*, an advanced measure of the mindreading ability for school age children, based on auditory stimuli. It would be expected that the *Voice Test* was related to other Theory of Mind measurements, in particular to those based on perceptive cues, and was not merely a measure of verbal ability (comprehension of the content of the spoken phrases and of the mental state terms). This hypothesis was confirmed by the pattern of correlations of the *Voice Test* with the other variable.

The *Voice Test* and the *Eyes Test* increased with age. As expected, the first-order false belief task was correctly solved by the school age sample, because children understand this type of false belief at around 4 years old (Wimmer & Perner, 1983; Wellman, Cross & Watson, 2001; Flynn, 2006). Most children also passed the second-order false belief task, in fact it is usually understood around 7 years old (Perner & Wimmer, 1985). There were a gender differences only in the second-order false belief task, as other ToM researches suggested (e.g., Charman, Ruffman & Clements, 2002).

It was found that the *Voice Test* performance correlated positively with other ToM tasks, both classical second-order false belief task and an advanced ToM task based on visual cues, the *Eyes Test*. Correlations remained significant also when verbal ability and age were controlled.

These data confirmed the convergent validity of the *Voice Test*: it assesses the ability to read mental states. Even though the correlations were high, but the correlation coefficient was not over 0.8 (that would have meant a quite perfect overlapping of the tests); this result can suggest that the construct of Theory of Mind is wide: it comprises
a lot of mental processes, both emotional and cognitive, and it is referred both to the “cognitive” understanding of false belief and to the “visual” and “auditory” communicative channels. However it is enough homogeneous because the different aspects composed Theory of Mind are strictly linked. This connection is more evident in daily life, during interpersonal relationships, where the mental state understanding from vocal cues is interlaced to that guided by visual cues and to the false belief comprehension.

The correlation between the Voice Test and a standardized verbal ability measure (PPVT – R), that follows closely the same trend of the other ToM tasks (False Belief task and Eyes Test), seems to confirm the close relationship and interdependence in development between ToM and language ability (e.g., De Villiers, 2000; Malle, 2002; Astington & Baird, 2005; Antonietti, Liverta Sempio & Marchetti, 2006; Siegal & Varley, 2006), even controlling for age (Milligan, Astington & Dack, 2007).

In conclusion, this study provided support for the construct validity of the Voice Test, a new measure of ToM based on perceptive (vocal) cues.

The previous and the present studies showed the psychometric properties of the test (reliability and validity); in the next section I present its standardization.

### 2.4. Phase 3: Standardization of the Voice Test

A standard test construction procedure requires – after the definition of the construct of interest, its operationalizations into measurable items and the analyses of the psychometric qualities (validity and reliability) – to standardize the observed scores,
indicating how they are distribute with regard to the specific population of people, through the $\zeta$-scores transformation.

The aim of the present study was to complete the *Voice Test* validation and standardization, identifying school age norms (standard scores for five group of age, from 6.5 to 11.4 years old).

### 2.4.1. Method

**Participants**

A total of 586 children participated in the study (the sample included the 220 children that were tested during the second phase of the research; see paragraph 2.3.). There were 295 males and 291 females aged between 78 and 137 months, attending six primary schools located in middle-class neighbourhood of Milan.

They had not psychological or neurological proclaimed pathologies and learning difficulties. A few participants were not Italian native speakers (less than 3%), but all were competent in speaking and understanding Italian. Their parents gave their informed consent to participate in the study.

Children were subdivided in five groups of age:

- 122 children (59 males and 63 females): 7 years old (78-89 months);
- 126 children (68 males and 58 females): 8 years old (90-101 months);
- 136 children (64 males and 72 females): 9 years old (102-113 months);
- 123 children (65 males and 58 females): 10 years old (114-125 months);
- 79 children (39 males and 40 females): 11 years old (126-137 months).
**Procedure**

Participants were tested in group (consisted of a maximum of 10 children; see paragraph 2.2.2.) in a quiet room at their school. Children were instructed to find and mark, in the answer sheets, the term, among the four mental words written within the same row, that better represented the speakers’ mental states (see specific instructions in Appendix).

Participants were said to remain silent and to pay attention to the sentence because they could listen to it once. Their comprehension of the test procedure and the appropriate level of the volume were tested using a trial item.

Before listening each item, were read the 4 options on the answer sheet and asked whether they did not understand any words or they were unsure of any words meaning; if they needed an explanation of a word, the experimenter read aloud some synonymous and an example from the Glossary (Appendix). The experimenter paused the audiotape between speech segments to give children the time they needed to choose the mental state term.

Then, at the end of the test, they received new instructions about the control task; the experimenter showed the answer sheet with two choice (“M” for male and “F” for female) for each item. Children listened again to the same items recorder on the CD (preceded by the trial item) and they marked the speaker’s gender in real time.

For each correct answer, children got one point, so the total score for both the *Voice Test* and the control test could vary from 0 to 20. The *Voice Test* took about 20 minutes.
2.4.2. Results

Descriptive statistics

Exploratory analysis suggested the Voice Test score to be normally distributed (Chart 2.8.): mean was 13.23; median and mode coincided and were 14,00; standard deviation was 3.32; skewness and kurtosis of the distribution were respectively -0.787 (s.d.: 0.10) and 0.363 (s.d.: 0.20). The distributions of the Voice Test in each group of age met normally requirements; Table 2.5. reports mean and standard deviation of the Voice Test in each group of age.

All the participants recognized correctly the speakers’ gender in the control task.

Chart 2.8. Voice Test total score distribution.
Table 2.5. Mean and standard deviation of the Voice Test total score in each group of age

<table>
<thead>
<tr>
<th>Age Group</th>
<th>Mean</th>
<th>s.d.</th>
</tr>
</thead>
<tbody>
<tr>
<td>78-89 months (7 years)</td>
<td>10.58</td>
<td>3.60</td>
</tr>
<tr>
<td>90-101 months (8 years)</td>
<td>12.12</td>
<td>3.03</td>
</tr>
<tr>
<td>102-113 months (9 years)</td>
<td>13.87</td>
<td>2.86</td>
</tr>
<tr>
<td>114-125 months (10 years)</td>
<td>14.99</td>
<td>1.87</td>
</tr>
<tr>
<td>126-137 months (11 years)</td>
<td>15.27</td>
<td>2.34</td>
</tr>
</tbody>
</table>

Influence of gender and age

Children’s gender was shown to have a significant effect upon children’s performance in the Voice Test: $t(584) = 2.437, p = 0.01$. Females had better performance ($M = 13.57$) than males ($M = 12.90$). The independent t test run for each group of age showed that gender difference began to become significant from 9 years old: $t(134) = 1.915, p = 0.05$; $t(121) = 2.518, p = 0.01$; $t(77) = 2.745, p < 0.01$, respectively for 9, 10 and 11 years old (Chart 2.9.).

In order to evaluate differences in the Voice Test among the different groups of age a one-way ANOVA was run with the five levels of age as the factor and the Voice Test total score as the dependent variable. This analysis revealed a significant effect of age on the performance in the test (Chart 12.10.), the Voice Test score increased with age: $F(4, 581) = 55.284, p < 0.001$.

Post-hoc test (Bonferroni t test) revealed significant differences among each group of age, except between 10 and 11 years old (Table 2.6.). There was not a significant interaction effect between age and gender on the Voice Test.
**Chart 2.9.** Gender differences in the Voice Test

**Chart 2.10.** Age differences in the Voice Test
Table 2.6. Post hoc analysis

<table>
<thead>
<tr>
<th>Group of age</th>
<th>Mean difference (s.d.)</th>
<th>Sig.</th>
</tr>
</thead>
<tbody>
<tr>
<td>7 years</td>
<td></td>
<td></td>
</tr>
<tr>
<td>8 years</td>
<td>- 1.54 (0.36)</td>
<td>p &lt; 0.001</td>
</tr>
<tr>
<td>9 years</td>
<td>- 3.29 (0.35)</td>
<td>p &lt; 0.001</td>
</tr>
<tr>
<td>10 years</td>
<td>- 4.41 (0.36)</td>
<td>p &lt; 0.001</td>
</tr>
<tr>
<td>11 years</td>
<td>- 4.68 (0.41)</td>
<td>p &lt; 0.001</td>
</tr>
<tr>
<td>8 years</td>
<td></td>
<td></td>
</tr>
<tr>
<td>7 years</td>
<td>1.54 (0.36)</td>
<td>p &lt; 0.001</td>
</tr>
<tr>
<td>9 years</td>
<td>- 1.75 (0.35)</td>
<td>p &lt; 0.001</td>
</tr>
<tr>
<td>10 years</td>
<td>- 2.87 (0.36)</td>
<td>p &lt; 0.001</td>
</tr>
<tr>
<td>11 years</td>
<td>- 3.15 (0.41)</td>
<td>p &lt; 0.001</td>
</tr>
<tr>
<td>9 years</td>
<td></td>
<td></td>
</tr>
<tr>
<td>7 years</td>
<td>3.29 (0.35)</td>
<td>p &lt; 0.001</td>
</tr>
<tr>
<td>8 years</td>
<td>1.75 (0.35)</td>
<td>p &lt; 0.001</td>
</tr>
<tr>
<td>10 years</td>
<td>- 1.12 (0.35)</td>
<td>p = 0.01</td>
</tr>
<tr>
<td>11 years</td>
<td>- 1.40 (0.40)</td>
<td>p &lt; 0.01</td>
</tr>
<tr>
<td>10 years</td>
<td></td>
<td></td>
</tr>
<tr>
<td>7 years</td>
<td>4.41 (0.36)</td>
<td>p &lt; 0.001</td>
</tr>
<tr>
<td>8 years</td>
<td>2.87 (0.36)</td>
<td>p &lt; 0.001</td>
</tr>
<tr>
<td>9 years</td>
<td>1.12 (0.35)</td>
<td>p = 0.01</td>
</tr>
<tr>
<td>11 years</td>
<td>0.27 (0.40)</td>
<td>n.s.</td>
</tr>
<tr>
<td>11 years</td>
<td></td>
<td></td>
</tr>
<tr>
<td>7 years</td>
<td>4.68 (0.41)</td>
<td>p &lt; 0.001</td>
</tr>
<tr>
<td>8 years</td>
<td>3.15 (0.41)</td>
<td>p &lt; 0.001</td>
</tr>
<tr>
<td>9 years</td>
<td>1.40 (0.40)</td>
<td>p &lt; 0.01</td>
</tr>
<tr>
<td>10 years</td>
<td>0.27 (0.41)</td>
<td>n.s.</td>
</tr>
</tbody>
</table>

Normative data

The Voice Test scores, for each group of age, were transformed in $z$-scores, defined by the formula: $(X - M) / \text{s.d.}^{14}$. The magnitude pf the $z$-score indicates how many standard deviations is the score away from the mean value and the sign (positive or negative) of the $z$-score shows whether the Voice Test score is above or below the mean. Tables 2.7., 2.8., 2.9., 2.10. and 2.11. show the correspondent $z$-score for each Voice Test score in the five groups of age.

---

$^{14} X$ is the Voice Test score, $M$ is the mean and s.d. is the standard deviation.
Table 2.7. Age 7 years (78-89 months): $\zeta$-scores

<table>
<thead>
<tr>
<th>Voice Test score ($X$)</th>
<th>$\zeta$</th>
<th>Cumulative %</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>-2,66</td>
<td>0,8</td>
</tr>
<tr>
<td>3</td>
<td>-2,11</td>
<td>1,6</td>
</tr>
<tr>
<td>4</td>
<td>-1,83</td>
<td>5,7</td>
</tr>
<tr>
<td>5</td>
<td>-1,55</td>
<td>9,8</td>
</tr>
<tr>
<td>6</td>
<td>-1,27</td>
<td>12,3</td>
</tr>
<tr>
<td>7</td>
<td>-0,99</td>
<td>19,7</td>
</tr>
<tr>
<td>8</td>
<td>-0,72</td>
<td>30,3</td>
</tr>
<tr>
<td>9</td>
<td>-0,44</td>
<td>41,8</td>
</tr>
<tr>
<td>10</td>
<td>-0,16</td>
<td>47,5</td>
</tr>
<tr>
<td>11</td>
<td>0,12</td>
<td>55,7</td>
</tr>
<tr>
<td>12</td>
<td>0,39</td>
<td>67,2</td>
</tr>
<tr>
<td>13</td>
<td>0,67</td>
<td>76,2</td>
</tr>
<tr>
<td>14</td>
<td>0,95</td>
<td>84,4</td>
</tr>
<tr>
<td>15</td>
<td>1,23</td>
<td>91,8</td>
</tr>
<tr>
<td>16</td>
<td>1,51</td>
<td>97,5</td>
</tr>
<tr>
<td>17</td>
<td>1,78</td>
<td>98,4</td>
</tr>
<tr>
<td>18</td>
<td>2,06</td>
<td>100,0</td>
</tr>
</tbody>
</table>

Table 2.8. Age 8 years (90-101 months): $\zeta$-scores

<table>
<thead>
<tr>
<th>Voice Test score ($X$)</th>
<th>$\zeta$</th>
<th>Cumulative %</th>
</tr>
</thead>
<tbody>
<tr>
<td>4</td>
<td>-2,68</td>
<td>1,6</td>
</tr>
<tr>
<td>5</td>
<td>-2,35</td>
<td>5,6</td>
</tr>
<tr>
<td>6</td>
<td>-2,02</td>
<td>7,1</td>
</tr>
<tr>
<td>7</td>
<td>-1,69</td>
<td>8,7</td>
</tr>
<tr>
<td>8</td>
<td>-1,36</td>
<td>11,9</td>
</tr>
<tr>
<td>9</td>
<td>-1,03</td>
<td>15,1</td>
</tr>
<tr>
<td>10</td>
<td>-0,70</td>
<td>26,2</td>
</tr>
<tr>
<td>11</td>
<td>-0,37</td>
<td>37,3</td>
</tr>
<tr>
<td>12</td>
<td>-0,04</td>
<td>46,8</td>
</tr>
<tr>
<td>13</td>
<td>0,29</td>
<td>62,7</td>
</tr>
<tr>
<td>14</td>
<td>0,62</td>
<td>81,0</td>
</tr>
<tr>
<td>15</td>
<td>0,95</td>
<td>88,9</td>
</tr>
<tr>
<td>16</td>
<td>1,28</td>
<td>96,0</td>
</tr>
<tr>
<td>17</td>
<td>1,61</td>
<td>99,2</td>
</tr>
<tr>
<td>18</td>
<td>1,94</td>
<td>100,0</td>
</tr>
</tbody>
</table>
Table 2.9. Age 9 years (102-113 months): \(\zeta\)-scores

<table>
<thead>
<tr>
<th>Voice Test score (X)</th>
<th>(\zeta)</th>
<th>Cumulative %</th>
</tr>
</thead>
<tbody>
<tr>
<td>6</td>
<td>-2,75</td>
<td>1,5</td>
</tr>
<tr>
<td>7</td>
<td>-2,40</td>
<td>2,2</td>
</tr>
<tr>
<td>8</td>
<td>-2,05</td>
<td>5,1</td>
</tr>
<tr>
<td>9</td>
<td>-1,70</td>
<td>7,4</td>
</tr>
<tr>
<td>10</td>
<td>-1,35</td>
<td>13,2</td>
</tr>
<tr>
<td>11</td>
<td>-1,00</td>
<td>20,6</td>
</tr>
<tr>
<td>12</td>
<td>-0,65</td>
<td>29,4</td>
</tr>
<tr>
<td>13</td>
<td>-0,30</td>
<td>39,7</td>
</tr>
<tr>
<td>14</td>
<td>0,05</td>
<td>54,4</td>
</tr>
<tr>
<td>15</td>
<td>0,40</td>
<td>70,6</td>
</tr>
<tr>
<td>16</td>
<td>0,74</td>
<td>80,1</td>
</tr>
<tr>
<td>17</td>
<td>1,09</td>
<td>91,2</td>
</tr>
<tr>
<td>18</td>
<td>1,44</td>
<td>97,8</td>
</tr>
<tr>
<td>19</td>
<td>1,79</td>
<td>100,0</td>
</tr>
</tbody>
</table>

Table 2.10. Age 10 years (114-125 months): \(\zeta\)-scores

<table>
<thead>
<tr>
<th>Voice Test score (X)</th>
<th>(\zeta)</th>
<th>Cumulative %</th>
</tr>
</thead>
<tbody>
<tr>
<td>10</td>
<td>-2,67</td>
<td>1,6</td>
</tr>
<tr>
<td>11</td>
<td>-2,13</td>
<td>3,3</td>
</tr>
<tr>
<td>12</td>
<td>-1,60</td>
<td>9,8</td>
</tr>
<tr>
<td>13</td>
<td>-1,06</td>
<td>19,5</td>
</tr>
<tr>
<td>14</td>
<td>-0,53</td>
<td>36,6</td>
</tr>
<tr>
<td>15</td>
<td>0,01</td>
<td>61,0</td>
</tr>
<tr>
<td>16</td>
<td>0,54</td>
<td>81,3</td>
</tr>
<tr>
<td>17</td>
<td>1,07</td>
<td>91,1</td>
</tr>
<tr>
<td>18</td>
<td>1,61</td>
<td>97,6</td>
</tr>
<tr>
<td>19</td>
<td>2,14</td>
<td>99,2</td>
</tr>
<tr>
<td>20</td>
<td>2,68</td>
<td>100,0</td>
</tr>
</tbody>
</table>
Table 2.11. Age 11 years (126-137 months): $z$-scores

<table>
<thead>
<tr>
<th>Voice Test score ($X$)</th>
<th>$z$</th>
<th>Cumulative %</th>
</tr>
</thead>
<tbody>
<tr>
<td>8</td>
<td>-3,11</td>
<td>3,8</td>
</tr>
<tr>
<td>10</td>
<td>-2,25</td>
<td>5,1</td>
</tr>
<tr>
<td>11</td>
<td>-1,82</td>
<td>6,3</td>
</tr>
<tr>
<td>12</td>
<td>-1,40</td>
<td>8,9</td>
</tr>
<tr>
<td>13</td>
<td>-0,97</td>
<td>13,9</td>
</tr>
<tr>
<td>14</td>
<td>-0,54</td>
<td>30,4</td>
</tr>
<tr>
<td>15</td>
<td>-0,12</td>
<td>50,6</td>
</tr>
<tr>
<td>16</td>
<td>0,31</td>
<td>68,4</td>
</tr>
<tr>
<td>17</td>
<td>0,74</td>
<td>88,6</td>
</tr>
<tr>
<td>18</td>
<td>1,17</td>
<td>94,9</td>
</tr>
<tr>
<td>19</td>
<td>1,59</td>
<td>98,7</td>
</tr>
<tr>
<td>20</td>
<td>2,02</td>
<td>100,0</td>
</tr>
</tbody>
</table>

2.4.3. Discussion and conclusion

The present study provided the *Voice Test* standardized scores for Italian school age children, from 78 to 137 months. Analyses showed that the *Voice Test* score increased with age and that it can discriminate well among age, in particular among seven, eight, nine and ten years old; while ten and eleven years old children performed fairly similarly.

The study confirmed that females are better than males in Theory of Mind (Charman, Ruffman & Clements, 2002; Baron-Cohen, 2003), in fact they had higher performance than males in the *Voice Test*. This gender difference begins to become evident and significant in older children, from nine years old.

The normative data of the *Voice Test* were calculated through the transformation in $z$-scores for each group of age: 6,5-7,4 years; 7,5-8,4 years; 8,5-9,4 years; 9,5-10,4 years; 10,5-11,4 years.
2.5. Further analysis on the Voice Test reliability

Reliability refers to the degree to which a test is consistent and stable in measuring what it is intended to measure; it is the consistency of a measuring instrument. A test is therefore reliable if it is consistent within itself and across time. In the paragraph 2.3, I analyzed, through the Cronbach’s alpha, a form of reliability called internal consistency reliability, that is the assessment of the consistency of results across items within the Voice Test.

In this paragraph I analyze further the Voice Test reliability, using the method of the test-retest. It consists in the administration of the Voice Test to the same sample in two different occasions, spaced out by a suitable period of time (not too much soon or later, because the retest could be distorted by a learning effect or by developmental change in children; Boncori, 2006). This type of reliability is very important especially for test, like the Voice Test, that assesses different aspects of a construct (Guilford, 1965), for example the Voice Test regards the comprehension of a wide range of complex mental states, both epistemic and emotional.

2.5.1. Method

Participants

Participants were 141 children (75 males and 65 females), including:

- 37 children (21 males, 16 females) aged 78-89 months (7 years old);
- 29 children (15 males, 14 females) aged 90-101 months (8 years old);
- 32 children (20 males, 12 females) aged 102-113 months (9 years old);
- 25 children (13 males, 12 females) aged 114-125 months (10 years old);
- 18 children (7 males, 11 females) aged 126-137 months (11 years old).

Children came from a middle-class background and attended the primary school in a town near Milan. Parents gave their written consent to allow children’s participation in this study.

**Measures**

Children were administered in group the *Voice Test*, in a quiet room, following the procedure described in the previous paragraph and in the Appendix.

In addition to the Voice Test, children were tested a standardized measure of receptive vocabulary, the *Peabody Picture Vocabulary Test – Revised* (PPVT-R) (Dunn & Dunn, 1981; Italian validation: Stella, Pizzoli & Tressoldi, 2000), already described in the paragraph 2.3.1. Non verbal intelligence quotient was also assessed using the *Raven Coloured Progressive Matrix* (RCPM) (Raven, 1984).

**Procedure**

The *Voice Test* was administered in group of 10 children. After four weeks (considered the better period between the two test sessions, in fact children unlikely could improve they mindreading ability in few weeks and they unlikely could remember the test, reducing a lot the learning effect) children were re-tested, exactly in the same way of the previous administration. Half the sample received the PPVT – R and the RCPM individually after the *Voice Test* administration in the test condition, and half received it individually in the re-test condition.
2.5.2. Results

Descriptive statistics

Table 2.12. and 2.13. show means and standard deviations of the Voice Test scores, both in the test and re-test conditions, in the entire sample and in each group of age. All children recognized correctly the speakers’ gender in the control test.

PPVT – R and RCPM scores were normally distributed and the had, respectively, a mean of 103,67 (s.d.: 14,35) and of 108,10 (s.d.: 14,55).

Verbal ability and IQ influences

PPVT – R correlated positively with the Voice Test both in test (Chart 2.11.) and re-test conditions ($r = 0,431$ and 0,467 respectively, $p < 0,001$), also controlling for age ($r = 0,342$ and 0,383 respectively, $p < 0,001$). RCPM did not correlate with the Voice Test (Chart 2.12.) in both conditions. A hierarchical regression analysis was conducted to assess the relative contribution of age, verbal ability and IQ to children’s performance in the Voice Test (test condition). The model explained the 35% of the variance (adjusted R squared: 0,351), $F (3, 137) = 25,255$ ($p < 0,001$). The equation was: Voice Test score = -1,53 + 0,44 age + 0,29 verbal ability + 0,025 IQ. Only the IQ independent variable did not achieve significance, while age and verbal ability were significant ($p < 0,001$).

Test-retest

Pearson’s product-moment correlation was calculated between the Voice Test scores obtained in the test and re-test conditions. It was 0,70 ($p < 0,001$).
It remained significant also controlling for verbal ability (PPVT – R) \( [r = 0.621, \ p < 0.001] \), for age \( [r = 0.570, \ p < 0.001] \) and for verbal ability and age \( [r = 0.505, \ p < 0.001] \).

**Table 2.12.** Descriptive statistics of the Voice Test in the test and re-test conditions.

<table>
<thead>
<tr>
<th></th>
<th>Mean</th>
<th>s.d.</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Test</strong></td>
<td>14.26</td>
<td>3.09</td>
</tr>
<tr>
<td><strong>Re-test</strong></td>
<td>15.25</td>
<td>3.15</td>
</tr>
</tbody>
</table>

**Table 2.13.** Descriptive statistics in the test and re-test conditions in each group of age.

<table>
<thead>
<tr>
<th>Age</th>
<th>Condition</th>
<th>Mean</th>
<th>s.d.</th>
</tr>
</thead>
<tbody>
<tr>
<td>7-year</td>
<td><strong>Test</strong></td>
<td>12.24</td>
<td>3.27</td>
</tr>
<tr>
<td></td>
<td><strong>Re-test</strong></td>
<td>12.89</td>
<td>3.44</td>
</tr>
<tr>
<td>8-year</td>
<td><strong>Test</strong></td>
<td>13.52</td>
<td>3.17</td>
</tr>
<tr>
<td></td>
<td><strong>Re-test</strong></td>
<td>14.55</td>
<td>2.81</td>
</tr>
<tr>
<td>9-year</td>
<td><strong>Test</strong></td>
<td>15.13</td>
<td>2.56</td>
</tr>
<tr>
<td></td>
<td><strong>Re-test</strong></td>
<td>16.06</td>
<td>2.27</td>
</tr>
<tr>
<td>10-year</td>
<td><strong>Test</strong></td>
<td>15.32</td>
<td>1.93</td>
</tr>
<tr>
<td></td>
<td><strong>Re-test</strong></td>
<td>16.60</td>
<td>2.16</td>
</tr>
<tr>
<td>11-year</td>
<td><strong>Test</strong></td>
<td>16.61</td>
<td>1.82</td>
</tr>
<tr>
<td></td>
<td><strong>Re-test</strong></td>
<td>17.89</td>
<td>1.78</td>
</tr>
</tbody>
</table>
Chart 2.11. Correlation between the Voice Test and the PPVT - R

Chart 2.12. Correlation between the Voice Test and the RCPM
2.5.3. Discussion and conclusion

The present study provided some further supports to the Voice Test reliability, through the test-retest method. The test score was highly correlated, in the same school age sample, to the test score after 4 weeks.

Moreover the research investigated the relationship between the Voice Test and two standardized measures of verbal (PPVT – R) and non-verbal (RCPM) ability. As in the previous study it was found that the Voice Test is strictly linked to the receptive language ability (paragraph 2.3.); the novel finding was that the non-verbal ability was not related to the performance in the Voice Test.

These data suggest that Theory of Mind is independent from the non-verbal intelligence quotient, while it is interdependent in development with verbal ability (e.g., Antonietti, Liverta Sempio & Marchetti, 2006; Milligan, Astington & Dack, 2007). The regression analysis confirmed this hypothesis, in fact only age and verbal ability were significant predictors of the Voice Test performance.

In conclusion, this research contributed to further validate the new Theory of Mind advanced instrument based on auditory cues created, the Italian Voice Test – child version, through the analysis of the test-retest reliability and the relationships with verbal and non verbal ability.
2.6. Conclusion: Is the Voice Test a suitable ToM measure for school age children?

The current research describes the three phases of the development, validation and standardization of a new Italian Theory of Mind advanced task, based on perceptive cues, for school age children (78-137 months): the Voice Test. The test, that can be administered in a group setting (ten children), consists of 20 auditory stimuli, that are spoken phrases with a neutral meaning. Children were asked to understand speakers’ complex mental states, choosing it among four mental words. In the control trial they were asked to recognize speakers’ gender. To summarize, the results indicate that the psychometric qualities (internal and test-retest reliability; construct validity) of the test are good and provide the normative data for a normal children population aged from 6,5 to 11,4 years.

In particular, the internal consistency is acceptable: Cronbach’s alpha is 0,69. This value can be influenced by the heterogeneity of mental states assessed, in fact the test regards both epistemic and complex emotional mental states. The test-retest reliability was high. The construct validity of the test was demonstrated through the positive correlations between the Voice Test and others classical (second-order False Belief task) and advanced (Eyes Test) ToM tasks. These results suggested that the construct of ToM is at the same time homogeneous, because of the strong correlations, and wide, because it comprises various types of mental states (i.e., epistemic states in the false belief task, epistemic and emotional states in the Eyes Test, epistemic and complex emotional states in the Voice Test) and different mental processes (i.e., different modalities used to recognize mental states, that use the cognitive, the visual or the auditory channel).
The *Voice Test* performance increase with age: significant differences were found among 7, 8, 9 and 10, while children aged 11 have a performance similar to those aged 10. Females obtained higher score in the test than males, as was found in other ToM tasks (Charman, Ruffman & Clements, 2002; Baron-Cohen, 2003).

The correlation between the *Voice Test* and standardized IQ measures showed that the test correlated with and was predicted by the receptive vocabulary ability (assessed with the PPVT – R), confirming results obtained with other classical and advanced ToM tests and the link between ToM and language ability (Astington & Baird, 2005; Antonietti, Liverta Sempio & Marchetti, 2006; Milligan, Astington & Dack, 2007).

Non verbal IQ, assessed with RCPM, did not correlate with the *Voice Test* showing its independence from intellectual ability, as found in other advanced ToM tasks based on perceptive cues (e.g., *Eyes Test* by Baron-Cohen et al., 1997, 2001a; *Voice Test* by Rutherford, Baron-Cohen and Wheelwright, 2001; *Reading the Mind in the Films* by Golan et al., 2006b; see paragraph 2.1.)

In conclusion the *Voice Test* can be considered a valid ToM measure, suitable for school age children. It will can be used to assess ToM in older children, besides other ToM tasks, in order to better grasp the complexity of mindreading ability during development. It will can be useful to study ToM in children with disabilities, such as children with autism or language impairment, but also blind children in the auditory communicative channel is very important because they cannot count on the visual one.

The *Voice Test* can also contribute to analyze individual differences in ToM, showing how and why different ToM abilities develop; for example mental state understanding from eyes, from voices and from the “cognitive” channel (e.g., knowledge of the situation) can differ both among children and within the children themselves.
Chapter 3

THEORY OF MIND AND EMOTIONAL DIFFICULTIES IN CHILDREN

Thanks to the ability to “read” one’s own and others’ mind, children can develop self-awareness (Howlin, Baron-Cohen & Hadwin, 1999), affect regulation abilities (Fonagy, Gergely, Jurist & Target, 2002) and social skills (Moore & Frye, 1991; Slaughter, Dennis & Pritchard, 2002; Astington, 2003; Cassidy et al., 2003) (see also chapter 1). Thus it can be hypothesized, on the basis of the influence that ToM has on children behavioral and emotional competences, that the mentalization ability have a protective function in children who are at risk of psychological problems (Fonagy & Target, 1996; 1998). Those who are not able (or not much able) to understand mental states (epistemic, emotional and motivational states; Astington, 2001) and to use them to give meaning (or the right meaning) to one’s own and others’ behaviors, can show social impairment, because they interact with other without understanding what the other have in his/her mind (Hughes & Leekam, 2004; Liverta Sempio & Marchetti, 2006).

To explain this idea it can be useful to refer to the severe pathologies within the autism spectrum conditions. Autism was the first psychopathology studied in the field of ToM: in 1985, Baron-Cohen, Leslie and Frith suggested that social and communicative impairment typical of persons with autism could be explained as a lack of mental states understanding ability. In fact many autistic individuals appear not able to understand
what other people have in their mind (plans, thoughts, beliefs, emotions…) and that the others can think differently than themselves; they often find the social environment unpredictable and incomprehensible and it is said that they seem to treat people and objects alike. These impairments can be on the basis of their problems in relating socially to others and cannot be explained referring uniquely to autistic’s mental retardation (because there are autistic children and adults with an intellectual quotient in the normal range and there are other individuals, such as those with Down’s syndrome, that are socially competent even if they are mentally retarded): the social difficulties in autism underline a difficulty in processing social information based on mental states understanding.

Researches found that autism is characterized by a severe delay in ToM precursors (e.g., sharing intentions, declarative pointing, shared attention, symbolic play) and poor performances in classical and advanced ToM tasks (e.g., Baron-Cohen, 1990; 1995, 1997, 2001; Happè, 1994, 1999; Buitelaar et al., 1999; Howlin, Baron-Cohen, Hadwin, 1999; Surian & Leslie, 1999; Frith, 2003; Fisher, Happè & Dunn, 2005; Hale & Tager-Flusberg, 2005; Perucchini, Muratori & Parrini, 2005; Colle, Baron-Cohen & Hill, 2007; Tager-Flusberg, 2007).

The present chapter has the aim – after reviewing literature about the link between ToM and social competence studied referring most of all to children with externalizing problems – to analyze deeper the relationship that could exist between mentalization and emotional difficulties, typical of internalizing behaviors.

In order to this aim, I present two researches. The first explored the relationship between ToM and psychological risk of conduct problems, somatization, depression and anxiety, finding a link between the poor mindreading ability and the risk to develop
depression and psychosomatic disorders. Starting from this result, the second research investigated in detail the relationship between ToM and somatic complaints in children. This work can contribute to the recent researchers’ interest in the study of children individual differences in social understanding (e.g., Repacholi & Slaughter, 2003; Hughes et al., 2005; Ronald et al., 2006).

3.1. Introduction: Theory of Mind point of view on children psychological problems

Developmental psychopathology studies traditionally distinguish between internalizing and externalizing problems (Mesman, Bongers & Koot, 2001). Internalizing problems include anxiety, somatization and depression; externalizing ones regard behavioral disorders such as antisocial behavior, conduct disorders and attention deficit-hyperactivity disorders. As Mesman, Bongers and Koot (2001) underlined, externalizing problems have generally received more research attention than internalizing ones and the same trend can be found also within ToM studies, as I explain in the next section. To study ToM in persons with psychological problems, characterized by difficulty in the socio-emotional functioning, can throw light upon the link between social and/or emotional abilities and mentalization development. So what about ToM in children with externalizing and internalizing problems? Quite few researches explored this subject; in the next two sections I outline the state of the art of ToM studies regarding children with behavioral and emotional problems.
3.1.1. Theory of Mind and externalizing problems

In these last years, ToM studies have moved their attention from severe psychopathological conditions (i.e. autism spectrum disorders), where mindreading disabilities seem to appear evident (Tager-Flusberg, 2007), to psychological conditions characterized by behavioral problems (reviews: Corcoran, 2000; Liverta Sempio, 2002; Sharp, 2006), searching data about the effect of ToM on social relationships. In other words, these researches have tried to answer to the question: is ToM necessary to social competence? Up till now it has not found a clear answer.

In fact, if on the one hand it appears evident ToM relevance for social and relational life (i.e., a good social adaptation is possible if we can recognize epistemic and emotional mental states, especially those in contrast with reality, such as distinguishing appearance from reality, realizing the existence of different emotions, desires and beliefs, understanding false beliefs), on the other hand this problem is not so easy to face (Liverta Sempio & Marchetti, 2005). It is not still clear the link between mentalization and social functioning (Repacholi & Slaughter, 2003) and it probably depends on definition of ToM adopted (Liverta Sempio & Marchetti, 2005).

Following a restricted definition, ToM (traditionally operationalized as a false belief comprehension) is sometimes a necessary, but not a sufficient condition for the social competence and probably other factors are necessary, such as affective involvement in the social situation, emotion (and not only epistemic) understanding and motivation (Aстington, 2003).

While using a wider definition (not restricted to false belief understanding), social behaviour is supposed based on a differentiation of mentalization abilities. Lucariello (2004; Lucariello et al., 2006) distinguished a “social” ToM and an “intrapersonal”
ToM, that develop separately. The social ToM, so called because we use it in social interactions, is the ability to understand others’ mental states and it is usually studied using the classical false belief task. Instead the intrapersonal ToM, that is on the basis of introspection, concerns the understanding of one’s own mental states and includes the ability to reflect, to maintain opposing objects or events representations and to learn. This view seems to be supported by neuroscientific studies with fMRI (Walter et al., 2004), but it can be suggested that, dividing self knowledge and knowledge of the other, it has the limit to not consider social behavior consisting of the understanding of both others and one’s own mental states; moreover it cannot explain how and when these two types of ToM integrate each others (Liverta Sempio, Marchetti, 2006).

Data from empirical researches on children with externalizing problems (such as conduct disorders, antisocial behaviors, bullying, attention deficit-hyperactivity disorders – ADHD) seem to convalidate Astington’s (2003) view of ToM as not strictly necessary, even if it is very important, for social competence.

One of the first studies regarding the link between behavioral problems and ToM was conducted by Happè and Frith (1996). They found that children with conduct problems are able to pass classical false-belief tasks, in spite of their lack of communicative and social skills, that would require mental states understanding, they experienced in everyday life.

Similarly, Hughes, White, Sharpen & Dunn (2000) found that “hard to manage” children, that showed antisocial, angry and unsympathetic behaviors, are able to solve false belief and affective perspective taking tasks.

Also children with disruptive behaviors studied by Sutton, Reeves and Keogh (2000) and with ADHD studied by Charman, Carrol and Sturge (2001) and by Perner, Kain and Barchfeld (2002) have not a mentalization deficit.
But a different situation is outlined in other studies: Hughes, Dunn and White (1998), found that “hard to manage” children had difficult in the affective perspective taking task; Buitelaar and colleagues (1999) and Liverta Sempio, Fabio & Tiezzi (2005) showed that ADHD children did not understand the second-order false beliefs.

In conclusion, researches failed to find a clear ToM deficit; following Happè and Frith (1996) these results can be explained hypothesizing that children with behavioural problems have a “theory of nasty mind”: they are able to read other’s mind, but they impute to the other bad intentions towards them, consequently they act aggressive behaviors. Such explanation (i.e., children with behavioral problems may show intact, but biased mindreading) guides ToM researchers to investigate not only the presence or absence of mindreading ability, but also its content, that is which type of mental states children attribute to their and others’ mind or, in other words, their mentalising style (Slaughter & Repacholi, 2003; Ronald et al., 2005; Sharp, 2006).

Recently Sharp, Croudace and Goodyer (2007) studied misinterpretation or biased mentalising in seven to eleven children, through the use of ambiguous peer-related social scenarios; they found that children with externalizing disorders have an overly positive style (i.e., unrealistic and overly positive self-perception).

A particular/skewed use of the mindreading ability could explain another behavioral problem: bullying. In fact bullies can solve false belief and emotion-false belief tasks (Sutton, Smith & Swettenham, 1999; Sutton, 2001, 2003) and it can be argue that they use their ToM ability for example in proactive aggression, in manipulating others’ mind, in excluding other children from social interactions and cause them distress, or when they organize a gang. So they have a good ToM, but they do not have positive social relationship: Sutton, Smith and Swettenham (1999) defined “macchiavellism” the
bullies’ attitude to be aware of others’ mental states (first of all emotions) and, at the same time, to be unable or unwilling to share mental states with others.

3.1.2. Theory of Mind and internalizing problems

While externalizing behaviors are ones in which the child is socially troublesome, and are characterized by verbal aggression, oppositional defiance and conduct problems, internalizing behaviors are ones in which the child is inwardly troublesome and are typified by social withdrawal, somatic complaints, loneliness, anxiety and depression. Internalizing problems therefore deal much more with emotional factors, as Kovacs and Devlin (1998, p. 47) wrote, they are “conditions whose central feature is disordered mood or emotion”.

These emotional disorders – depression, anxiety, somatization – lead inevitably to social and interpersonal difficulties and can interfere with cognitive processing. Specifically, children with depression have low self-esteem and high level of sadness (e.g., Blumberg & Izard, 1985) and suffer poor peer relationships and self-isolation (e.g., Renouf, Kovacs & Mukerji, 1997). They also have some cognitive distortions and misinterpret external stimuli, for example it was seen that they are not able to properly and accurately decode nonverbal information in communication (pitch, tempo, voice inflection) (Emerson, Harrison & Everhart, 1999; Segrin, 2000).

Children with anxiety disorders feel much more fear, worry and apprehension (e.g., Brady & Kendall, 1992) and they are less assertive, shyer and more withdrawn than non anxious children (Ginsburg, La Greca & Silverman, 1998).

Somatization, that is the tendency to experience somatic distress and symptoms, not clearly explicable in terms of a precise medical diagnosis or disorder, and to attribute
them to physical illness (Lipowski, 1988), is characterized by the misattribution of the emotions, considered only as physical sensations and not also as mental states (in fact emotions are physical experiences as well as psychological or cognitive ones; e.g., Damasio, 1999). The somatic symptoms in children often interfere with school, home life and peer relationships (Walker, Garber & Greene, 1994; Garralda, 1999).

Thus the difficulties typical of internalizing problems in the cognitive processing could suggest that also the mindreading ability is impaired. But I do not find researches regarding ToM in children with internalizing problems. This field of study could be fruitful, because a specific mentalising bias can characterize these children and can be linked with their difficulties in cognitive, emotional and social life.

Studies on adults with depression (Inoue et al., 2004; Lee et al., 2005; Harkness et al., 2005; Inoue, Yamada & Kanba, 2006) found that their difficulties in social functioning can be caused by alterations in ToM and, more specifically, that they seem to make not accurate judgments of visual cues about what another person is thinking or feeling.


About children population, Sharp, Croudace and Goodyer (2007), in the research cited above (paragraph 3.1.1.), studied also ToM in children with emotional problems, but they failed to find a specific relationship between mentalising styles and symptoms of emotional disorders. This result was probably due to the type of ToM task used, that was aimed to study specific mentalising styles during peer interaction.

Recently Muris, Mayer, Vermeulen and Hiemstra (2007) studied children’s perception and interpretation of anxiety-related physical symptoms in children aged 4-13 years old and found that from 7 years children are more able to link physical symptoms to anxiety and that ToM abilities (assessed using the TOM Test by Muris et al., 1999, a
narrative task; see paragraph 2.1.) appear to have a significant impact in the interpretation of these symptoms and in emotional reasoning.

There are not other studies on this matter; the present work has the aim to begin to fill the gap regarding internalization problems in ToM studies.

3.2. Study 1: Theory of Mind and psychological risk

Till now, ToM researches have focused their attention on children's externalizing problems, in order to find if behavioral difficulties are related with and influenced by mentalizing ability. As I reviewed in the previous paragraph, there are not clear results on this subject, probably due for example to the various psychological problems analyzed and especially to the different ToM definitions adopted (i.e., ToM regarding the presence or absence of the ability to understand cognitive mental states or both cognitive and emotional mental states or ToM as a “style”) and measurement used to assess it.

The present study wants to widen ToM perspective in three directions. First, it analyzes not only behavioral problems, but also emotional ones (i.e. internalizing problems: anxiety, depression and somatization), because ToM deals both with social functioning and emotional functioning (see chapter 1). So this research contributes to the debate regarding externalizing problems and mentalization and tests the hypothesis that difficulty in ToM is linked to the emotional difficulties typical of internalizing behaviors.
Second, the study has the aim to analyze ToM not in children with a diagnosed psychopathology (e.g., ADHD, conduct disorder, autism, depression), but in children who can risk to develop a psychopathology, specifically at risk of conduct disorders, anxiety, depression and somatization. This is important because on the one hand it can account for individual differences in mentalization development, on the other hand it can be useful to early mental health problems identification in the school context (Mass Levitt et al., 2007).

Third, it focuses on children attending primary school (while the attention is usually on preschool children externalizing problems; e.g. Hughes et al., 2000; Charman, Carroll, Sturge, 2001; Perner, Kain, Barchfeld, 2002), assessing mentalization ability (that includes also emotion understanding) through not only traditional false belief tasks, but also an advanced and ecological ToM tasks based on vocal cues (i.e. the Voice Test, described in chapter 2).

The major goal of this study is to investigate the relationship between ToM and psychological risk, in particular the risk to develop externalizing (conduct disorders) and internalizing (depression, anxiety and somatization) disorders, in school age children.

3.2.1. Method

Participants
A sample of 112 children (56 males and 56 females), aged between 86 and 122 months (mean age: 103,49 months; standard deviation: 11,03), attending the primary school in two towns near Milan. Children were all Italian native speakers and came from the middle-class background (as assessed on father’s profession); they had not
psychological or neurological proclaimed pathologies and learning difficulties. Participant’s parents gave their written consent to allow children’s participation in the study.

Participants were subdivided in two groups of age:

1. 55 children (26 males and 29 females) aged between 86 and 104 months (M: 93,76 months; s.d.: 5,19);
2. 57 children (30 males and 27 females) aged between 105 and 122 months (M: 112,88 months; s.d.: 5,71).

Measures

The measures consisted of three ToM tasks, a questionnaire regarding the psychosocial risk and a verbal ability test.

Children were tested on the classical first-order (“The deceptive box”) (Perner, Leekam & Wimmer, 1987) and second-order false belief tasks (“Look prediction”) (Sullivan, Zaitchik, & Tager-Flusberg, 1994; Italian version: Antonietti et al., 1999), already described in chapter 2.3.1. The complex mental states (epistemic and emotional states) understanding was also assessed using the Voice Test, a new advanced ToM instrument, based on 20 vocal cues, for children aged 6,5-11,4 years (see chapter 2 and Appendix).

The psychosocial risk in children was assessed using a self-report measure, administered in a group setting: the Seattle Personality Inventory – Revised (S.P.I. – R) (Greenberg, 1994), adapted and validated on an Italian school age sample by Tani and Schneider (1998). Children reported of their own behavioural symptomatology, focuses on feelings and behaviours, on a four-point responses scale (from 1 “almost never” to 4 “almost always”).
It contains 32 items subdivided in four subscales focused on four different symptomatology dimensions (plus a control Lie Scale): conduct problems (8 items; e.g., “Do you talk in class a lot when you are not supposed to?”), anxiety (6 items; e.g., “Are you afraid to try new things?”), somatization (5 items; e.g., “Do you get a lot of pains in your body?”) and depressive symptoms (5 items; e.g., “Do you feel like crying a lot of the time?”).

A test of verbal mental age was individually administered: the *Peabody Picture Vocabulary Test – Revised* (PPVT – R) (Dunn & Dunn, 1981; Italian standardization: Stella, Pizzoli & Tressoldi, 2000).

**Procedure**

Children were tested in their school, in a quiet room. Each child participated in two testing sessions, conducted 10 days apart. In the first session children completed, individually, the PPVT – R and the first order false belief task (1\textsuperscript{st} FB). In the second session the second order false belief task (2\textsuperscript{nd} FB) was administered individually, then the S.P.I. – R and the *Voice Test* were administered in group of ten children.

**3.2.2. Results**

Table 3.1. shows means and standard deviations for all variables in the total sample and within the two groups of age.

Exploratory analyses suggested all variables to be normally distributed, except the two False Belief tasks: one children (one for each group of age) failed the 1\textsuperscript{st} FB task and 17 children (13 within the younger children group, that is the group 1) did not solved
correctly the 2\textsuperscript{nd} FB task, as expected. The 1\textsuperscript{st} FB task was excluded from analyses; nonparametric test were used with the 2\textsuperscript{nd} FB results.

All children answered correctly to the control questions of the False Belief tasks and recognized correctly the persons’ gender in the Voice Test control task.

The S.P.I. – R Lie scale did not correlate with the subscales of the risk of depression, anxiety and somatization (respectively: \( r = -0,18; -0,05; \) and \(-0,05\)) and correlate negatively with the subscale of the risk of conduct disorders \( (r = -0,493, p < 0,001) \). So I will consider the likely influence of the social desiderability in children with an high score in the subscale of the risk of conduct disorders when I will interpret the data.

<table>
<thead>
<tr>
<th></th>
<th>Total sample</th>
<th>1</th>
<th>2</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>N=112</td>
<td>N=55</td>
<td>N=57</td>
</tr>
<tr>
<td>M</td>
<td>s.d.</td>
<td>M</td>
<td>s.d.</td>
</tr>
<tr>
<td><strong>Age</strong></td>
<td>103,49</td>
<td>11,03</td>
<td>93,76</td>
</tr>
<tr>
<td><strong>PPVT – R</strong></td>
<td>103,63</td>
<td>12,61</td>
<td>102,87</td>
</tr>
<tr>
<td><strong>Voice Test (0-20)</strong></td>
<td>13,54</td>
<td>2,49</td>
<td>12,93</td>
</tr>
<tr>
<td><strong>Anxiety (0-4)</strong></td>
<td>2,09</td>
<td>0,56</td>
<td>2,10</td>
</tr>
<tr>
<td><strong>Depression (0-4)</strong></td>
<td>1,96</td>
<td>0,43</td>
<td>1,95</td>
</tr>
<tr>
<td><strong>Conduct Probl. (0-4)</strong></td>
<td>1,65</td>
<td>0,43</td>
<td>1,65</td>
</tr>
<tr>
<td><strong>Somatization (0-4)</strong></td>
<td>1,92</td>
<td>0,53</td>
<td>2,05</td>
</tr>
</tbody>
</table>
Influence of gender, age and verbal ability

The independent sample $t$ test did not find significant differences between males and females scores in all variables. The Mann-Whitney $U$ non parametric test found that females had better performance in the 2nd FB task (only 4 failed; mean rank: 61) than males (13 failed; mean rank: 52): $U (110) = 1316,00, p = 0,018$.

The PPVT – R and the S.P.I. – R subscales did not differ between the two group of age, excepted for the subscale of the risk of somatization, that decreased with age: $F (1,110) = 6,561, p = 0,012 (\eta^2 = 0,056)$.

The Voice Test performance increased with age: $F (1,110) = 6,964, p = 0,010$. Scores in this test follow the normative data presented in the second chapter.

The bivariate correlations between PPVT – R and other variables did not find any significant correlation of verbal ability with psychosocial risk and ToM.

Correlations

In order to investigate the link between ToM and psychosocial risk, bivariate correlations between the Voice Test and the S.P.I. – R were computed (Table 3.2): the Voice Test performance correlated negatively with the risk of depression and somatization subscales.

The nonparametric correlation Spearman’s rho showed significant correlations between the 2nd FB and the Voice Test ($rho = 0,189, p = 0,045$) and the 2nd FB and the subscale of the risk of depression ($rho = - 0,236, p = 0,012$).

Both correlations between the Voice Test and the risk of depression and the Voice Test and the risk of somatization (Charts 3.1. and 3.2.) remained significant also controlling for age: respectively $r = - 0,235, p = 0,013$ and $r = - 0,216, p = 0,022$. 
The partial correlations (controlling for age) among the four S.P.I.-R subscales remained significant, in particular: anxiety and depression $r = 0.421$ ($p < 0.001$); anxiety and somatization $r = 0.229$ ($p = 0.016$); depression and conduct problems $r = 0.265$ ($p = 0.005$); depression and somatization $r = 0.412$ ($p < 0.001$); conduct problems and somatization $r = 0.226$ ($p = 0.017$).

**Table 3.2.** Bivariate correlations between ToM and psychosocial risk

<table>
<thead>
<tr>
<th></th>
<th>Anxiety</th>
<th>Depression</th>
<th>Conduct Problems</th>
<th>Somatization</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Voice Test</strong></td>
<td>n.s.</td>
<td>$r = 0.216$</td>
<td>n.s.</td>
<td>$r = 0.282$</td>
</tr>
<tr>
<td></td>
<td></td>
<td>$p = 0.022$</td>
<td></td>
<td>$p = 0.003$</td>
</tr>
<tr>
<td><strong>Anxiety</strong></td>
<td>1</td>
<td>$r = 0.422$</td>
<td>n.s.</td>
<td>$r = 0.209$</td>
</tr>
<tr>
<td></td>
<td></td>
<td>$p &lt; 0.001$</td>
<td></td>
<td>$p = 0.027$</td>
</tr>
<tr>
<td><strong>Depression</strong></td>
<td></td>
<td>1</td>
<td>$r = 0.266$</td>
<td>$r = 0.390$</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>$p = 0.005$</td>
<td>$p &lt; 0.001$</td>
</tr>
<tr>
<td><strong>Conduct Problems</strong></td>
<td></td>
<td></td>
<td>1</td>
<td>$r = 0.199$</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>$p = 0.035$</td>
</tr>
</tbody>
</table>

**Chart 3.1.** Correlations between the Voice Test and the risk of depression
3.2.3. Discussion

The present study had the aim to explore the relationship between the risk of psychological problems and Theory of Mind in school age children.

ToM was assessed using both traditional and advanced tasks. As expected most of the participants passed the classical False Belief tasks, in fact the first order false belief is usually understood around 4 years (Wellman, Cross & Watson, 2001) and the second order false belief is understood at age 7 (Perner & Wimmer, 1985).

It was found that female had better performance in the second order false belief task than male, confirming the trend found by Charman, Ruffman and Clements (2002) in the first order false belief performance. The advanced ToM instrument used in this study was the Voice Test, that assesses the comprehension of complex epistemic and
emotional mental states from vocal cues; as showed in the second chapter the
performance in this test increase with age and scores follow normative data.
The psychological risk was assessed using the *Seattle Personality Inventory – Revised*
(Greenberg, 1994; Tani & Schneider, 1998). In particular it regards two types of
psychological risk: the risk to develop externalizing behaviors (i.e. conduct problems)
and internalizing behaviors (i.e. depression, anxiety and somatization). Younger
children (group 1) that participated to this study showed higher level in the risk of
somatization than older children (group 2). On the contrary, the literature on
somatization revealed that somatic complaints tend to increase with age, from
childhood to adolescence (Garber, Zeman, Walker, 1990; Perquin et al., 2000). The
result found in this study can be due to the specific instrument used to assess the risk
of somatization. In fact this S.P.I. – R subscale consists of five somatic symptoms
(headache, stomach ache, pain, nightmares and retching) usually complaint by younger
school age children, while it did not assess other symptoms such as tiredness, weakness,
pains in the joints, backache or multiple pains (e.g. simultaneously head ache and
stomach ache), that are experienced in older children (preadolescent) (Perquin et al.,
2000). It can be suggested that in this study the older children did not found in the
questionnaire those somatic symptoms they complaint.

The inter-correlations among the four subscale (conduct disorders, anxiety, depression
and somatization) of the S.P.I. – R are similar to those obtained by Tani and Schneider
(1998), except for the relation between anxiety and conduct disorders: they found a
weak positive correlation ($r = 0.18, p<0.005$), while in this study they do not correlate
significantly. The correlations among these four dimensions, as also noted Tani and
Schneider (1998, p. 522), confirm psychiatric literature on comorbidity (American
Psychiatric Association, 2000), especially that referring to internalizing problems (e.g.

The research did not find a relationship between ToM and the subscale of the risk of conduct problems. This result could be altered, considering that children who had higher score in the subscale showed higher social desiderability (low score in the lie subscale), nevertheless it seems to confirm Hughes, White, Sharpen and Dunn (2000), Sutton, Reeves and Keogh (2000), Charman, Carroll and Sturge (2001), Perner, Kain and Barchfeld (2002) results. These researchers found that the performance in ToM tasks in children “hard to manage”, with disruptive behavior and ADHD were not different from that of children without externalizing behaviors (see paragraph 3.1.1.).

The present study did not also find a correlation between ToM and risk of anxiety in children. There are not research on this matter; future studies should investigate deeper this relationship using specific instruments to assess anxiety and comparing children with a diagnosis of anxiety disorder and children without these symptoms.

The new relationships found in this research were those between ToM and the risk to develop two internalizing behaviors: depression and somatization. These correlations remained significant also controlling for age.

The results suggest that children that are not able to recognize correctly epistemic and emotional mental states from vocal cues are those who develop more depressive symptoms (e.g., to be unhappy, to cry, to suffer from lack of appetite and insomnia) and more psychosomatic symptoms (e.g., headache and stomach ache).

The link between ToM and depression was found in adults (e.g., Harkness et al., 2005; Inoue, Yamada & Kanba, 2006), using tasks based on visual cues. In the present study this link was extended even to child at risk of this psychopathology. Moreover, because the link was between the Voice Test, a ToM task based on vocal cues, and the risk of
depression, the research confirms the misinterpretation of auditory stimuli (i.e. lack of accuracy in decoding nonverbal cues in communication) found by Emerson, Harrison and Everhart (1999) and Segrin (2000) in children with depression.

Finally, the relationship between ToM and somatization were never considered in ToM studies. It is known that persons who experienced somatic symptoms, not explained by a medical diagnosis, have poor emotion awareness and often fail to identify emotions and to differentiate them from physiological states (Waller & Scheidt, 2006). Some researches explored the link between emotion understanding and somatization in school age children.

Rieffe, Meerum Terwogt and Bosch (2004) failed to found a poor emotion comprehension in children (aged 8-12 years) that experienced a lot of somatic symptoms, but they showed that this group of children attributed negative emotion (and, in particular, more fear than anger) to the characters in stories regarding peer relationship situations.

Jellesma, Rieffe, Meerum Terwogt and Kneepkens (2006) pointed out that the clinical group of children (8-13 years) with functional abdominal complaints and normal children reporting many somatic complaints shared the presence of negative moods, symptoms of depression, poor sense of coherence and two aspects of the emotion awareness: difficulty in emotion differentiation and communication. The two groups differed only in the awareness of the bodily sensations of emotions: children with many somatic complaints are strongly focused on bodily sensations of emotions, attributing the physiological phenomena that accompany emotions to an organic problems.

Recently also Rieffe and colleagues (2007) found that children (10-16 years old) that report many somatic complaints are focused on bodily sensations, moreover they are not able to differentiate and identify correctly the emotions.
In the present study it was found a link between the risk of somatization and the difficulty to recognize not only emotional, but also epistemic mental states from vocal cues. The researches I cited above, regarding somatization and emotion understanding, used narrative tasks; I do not know researches that investigate in children with psychosomatic problems the comprehension of nonverbal cues in communication. Because children often externalize their psychological problems through somatic symptoms (Garralda, 1999) and ToM regards the understanding of psychological states, it can be argued that children who are less able to mindread found also hard to elaborate mentally their problems and cannot read their emotions from bodily sensation (physiological arousal) of emotions. So on the one hand they express inner problems through the body, without mentalize them. On the other hand they do not recognize mental states based on body signals; the vocal cues, that are nonverbal aspects of mental states, can be considered, in a wide sense, as bodily signals (for example, the fear can produce a faltering voice or the muscular tension, due to anger, can cause a shrill voice), like visual cues, that this research did not considered.

The present study contributed to develop a deeper comprehension of children at risk of depression and somatization: it suggests a new variable, the mentalization ability, that could be considered in children’s internalizing behaviors, besides affective and personality variables (e.g., Eminson, 2001; Van Leeuwen, Mervielde, Braet, 2004; Waller, Scheidt, 2006; Nelson et al., 2007).
3.3. Study 2: Theory of Mind and somatic complaints

In order to analyze deeper the relationship between ToM and the risk of somatization in school age children, found in the previous research, it was developed another study. The choice to study in detail how ToM and somatization are linked proceeds from two reasons.

First, the presence of physical symptoms or painful complaints of unknown aetiology is a common occurrence in children population (Garber, Walker & Zeman, 1991; Campo & Fritsh, 1994; Perquin et al., 2000): about 30% of American children complain headache, stomach ache, tiredness and other symptoms once a week and only few complaints are due to a medical reason (Campo et al., 1999).

Normal children can express emotional distress in terms of somatic complaints, that are usually transitory and do not affect children’s functioning. But what about when children, from a not clinical population, experienced often and for a long time these symptoms? This problem seems increased in these last years (Santalahti et al., 2005) and it requires urgent investigations on its aetiology and on factors that could be associated to somatic complaints.

Somatization aetiology appears very complex, because includes biological, psychological and social factors. Some researchers underlined that psychological causes are not necessary to develop this problem (Walker & Garber, 2003), nevertheless there are some psychological features that can be found in children with somatization (Garralda, 1999): the presence of a time relationship between a likely stress and physical symptoms; the severity of handicap from the symptom that is out of keeping with the
established pathophysiology; concurrent psychiatric disorder\textsuperscript{15} (in particular anxiety and depression; e.g., Garber, Walker & Zeman, 1991; Campo & Fritsch, 1994; Hofflich, Hughes & Kendall, 2006); the presence of specific characteristics of the family and illness factors (e.g., parents’ physical and/or psychological illness, their worry of illness, the inclination to dissociate bodily and psychological experiences, the lack of communication on emotions, the degree and nature of the parental concerns about the child’s symptoms; Walker & Greene, 1989; Walker, Garber & Greene, 1993; Garralda, 1996; Eminson, 2001).

Moreover it was found that the mother-child attachment relationship plays a role in the development of somatic symptoms; in particular those who have an insecure attachment experience higher level of somatization (Hagekull & Boholin, 2004; Waller & Scheidt, 2004).

ToM could be a variable associated with somatic complaints in children, directly (as the previous study seems to suggest) and indirectly, considering the importance of the family contexts variables, such as attachment (i.e. insecure attachment is related to poor performance in ToM tasks) (Fonagy, Redfern & Charman, 1997; Fonagy & Target, 2001; Fonagy, Gergely & Target, 2007) and maternal mindmindedness (Meins et al., 2001) for ToM development (Liverta Sempio, 2002; Lecciso, 2005; see also chapter 1).

The second reason on the basis of this study is the close relationship that researches found between somatization and emotion. On the one hand both are in deal with the

\textsuperscript{15} Most children with somatization disorders do not have psychiatric comorbidity but a third to a half do (Garralda, 1999). Zwaigenbaum and colleagues (1999) found that somatic complaints at age 13-16 predict psychopathology 4 years later; Hotopf and colleagues (1998) showed that abdominal pains, not due to medical causes, during childhood (assessed at 7, 11 and 15 years old) persist also in adulthood and predict psychiatric disturbs at 36 years old. Nevertheless longitudinal researches on this subject are few and the causal relationship between somatization and following psychiatric disorders should be analyze deeper (Fritz, Fritsch & Hagino, 1997).
body. In fact somatization is characterized by persistently complaints of varied physical symptoms that have no identifiable physical origin; emotions can be considered, following Damasio’s view (1995, 1999), brain representations of body states: the brain links the body change with the emotion that accompanies it.

On the other hand somatization can be considered as an emotion understanding difficulty. Few researches on children population investigated the link between somatic complaints and emotions. I have already presented in the previous paragraph (3.2.3) the studies developed by Rieffe, Meerum Terwogt and Bosch (2004), Jellesma and colleagues (2006) and Rieffe and colleagues (2007) on this subject, that underlined the difficulty in emotion understanding in children with frequent somatic complaints. Another research, by Meerum Terwogt and colleagues (2006), found that some emotional variables (e.g., negative emotions, problems in handling negative emotions, coping response) and self-esteem predict psychosomatic symptoms in school age children (8-13 years). Recently, Jellesma, Rieffe and Meerum Terwogt (2007) showed that somatic complaint are strongly related to negative moods.

In conclusion, using words typical of ToM studies, these researches on emotions underlined the deficit, in persons’ with somatization problems, to consider emotions as mental states (psychological level, that is connected with bodily sensations of emotions), to understand them (e.g., difficulty in emotion identification and understanding, emotion misattribution) and to use mental state terms to express emotions. These difficulties would affect the focalization on somatic cues.

The present study has the aim to investigate, I think for the first time within ToM studies, the relationship between ToM and somatic complaints in a normal school age children population.
Starting from the result found in the first study (par. 3.2.), regarding the link between the poor ability in recognizing complex mental states from vocal cues and the risk of somatization, this research wants to better understand this link, using a specific instrument to assess somatic complaints and analyzing not only the mental states understanding from vocal cues but also from visual cues.

3.3.1. Method

Participants

96 children (47 males and 49 females), aged between 84 and 132 months (mean age: 107.91 months; standard deviation: 14.04) participated in the study. They attended the primary school in a town near Milan and were all Italian native speakers. Prior to conducting the research, permission to participate was obtained from parents. They were from middle-class background (as assessed from fathers’ job).

The sample was subdivided in 2 groups:

1) 47 children (22 males and 25 females), aged between 84-107 months;

2) 49 children (25 males and 24 females), aged between 108-132 months.

Measures

Theory of Mind was assessed with the classical first (“The deceptive box”; Perner, Leekam & Wimmer, 1987; Italian adaptation: Liverta Sempio & Marchetti, 2001) and second order (“Look prediction”; Sullivan, Zaitchick & Tager-Flusberg, 1994; Italian adaptation: Antonietti et al., 1999) false belief tasks, already used in the first study (and described in chapter 2.3.1).
Children were also administered two advanced ToM tasks based on perceptive cues: the *Voice Test* (see chapter 2 and Appendix), that assesses the ability to recognize complex mental states from voices, and the *Eyes Test* – Child version (Baron Cohen et al., 2001b; Italian translation: Livertu Sempio, Marchetti & Castelli, 2003), that assesses epistemic and emotional states understanding from visual cues (eyes expressions) (it is described in detail in paragraph 2.1.1.).

Somatic complaints were assessed by the *Somatic Complaint List* (SCL) (Rieffe, Meerum Terwogt & Bosch, 2004). The questionnaire consists of a list of 8 items: 6 somatic complaints that are common in younger and older children (e.g., have an headache, a stomach ache, to be tired, weak) and 2 items referred to “feel fine” and “well”. Children were asked to fill out each item on a likert-type scale (from 1 “never” to 3 “often”).

Finally children verbal and non verbal IQs were tested. Verbal IQ was assessed with the *Peabody Picture Vocabulary Test – Revised* (PPVT – R) (Dunn & Dunn, 1981; Italian standardization: Stella, Pizzoli & Tressoldi, 2000), a measure of the receptive vocabulary. The non verbal IQ was assessed with the *Raven Coloured Progressive Matrix* (RCPM) (Raven, 1984).

**Procedure**

Children were tested in two sessions, in a quiet room within their school, conducted 2 weeks apart.

In the first session their nonverbal and verbal IQs were assessed individually, then they received the first order false belief task. In the second session they completed in group of 10 children the *Voice Test* and the *Somatic Complaints List*, then they tested individually with the *Eyes Test* and the second order false belief task.
3.3.2. Results

Descriptive statistics

The variables were normally distributed, excepting the first order false belief task (only 1 child failed) and the second order false belief task (10 children in the group 1 and 4 in the group 2 did not pass the task).

All children answered correctly to false belief tasks control questions and all children passed the Voice Test control task (gender recognition). All children recognized correctly the gender of 23 of 28 persons at least in the Eyes Test (53,1% recognize all the persons’ gender correctly).

Table 3.4. shows means and standard deviations for normally distributed variables, in the entire sample and in each group of age.

<table>
<thead>
<tr>
<th></th>
<th>Total N=96</th>
<th>Group 1 N=47</th>
<th>Group 2 N=49</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age (months)</td>
<td>M</td>
<td>s.d.</td>
<td>M</td>
</tr>
<tr>
<td></td>
<td>107,91</td>
<td>14,04</td>
<td>95,66</td>
</tr>
<tr>
<td>PPVT – R</td>
<td>104,28</td>
<td>10,11</td>
<td>102,45</td>
</tr>
<tr>
<td>Raven (IQ)</td>
<td>107,81</td>
<td>14,50</td>
<td>108,40</td>
</tr>
<tr>
<td>Voice Test (0-20)</td>
<td>15,01</td>
<td>2,36</td>
<td>13,87</td>
</tr>
<tr>
<td>Eyes Test (0-28)</td>
<td>18,31</td>
<td>3,00</td>
<td>17,26</td>
</tr>
<tr>
<td>SCL (8-24)</td>
<td>15,40</td>
<td>1,84</td>
<td>15,47</td>
</tr>
</tbody>
</table>
**Gender, age and IQ influences**

The independent sample t test found the female (M = 15,65) had a better performance in the *Voice Test* than males (M = 14,34):  \( t(94) = 2,823, p=0,006 \) (Chart 3.3.).

There were not gender differences in the other variables.

As expected, performance in advanced ToM tasks increase with age:  \( F(1, 94) = 27,362, (\eta^2 = 0,225), p < 0,001; \) and  \( F(1, 94) = 12,801, (\eta^2 = 0,120), p = 0,001, \) respectively for the *Voice Test* (in line with the normative data presented in the second chapter) and the *Eyes Test*.

No age influence was found in the other variables.

Non verbal IQ (RCPM) correlated with PPVT – R (\( r = 0,323, p = 0,001 \)), but it did not correlate the other variables. Verbal ability correlated positively only with the *Voice Test* score:  \( r = 0,221, p = 0,031 \).

**Chart 3.3.** Gender and age differences in the Voice Test performance
Bivariate correlations

The Eyes Test and the Voice Test correlated positively: \( r = 0.310, p = 0.002 \).

The SCL score correlated negatively only with the Voice Test: \( r = -0.236, p = 0.020 \).

Partial correlation, controlling both for age and PPVT – R, was significant: \( r = -0.219, p = 0.34 \).

3.3.3. Discussion

The present research analyzed the relationship between ToM and somatic complaints in school age children. It confirmed the correlation between higher level of somatic complaints, that characterized psychosomatic problems, and the poor performance in the Voice Test, even controlling for age and verbal ability: children who experienced more somatic symptoms have difficulty to correctly recognize complex epistemic and emotional states from voices.

About age, gender and verbal ability influences on the Voice Test, I refer to the second chapter, where it was explained that this advanced ToM task increases with age, is dependent from the receptive language ability (but not from intellectual abilities) and females has higher score than males, as other ToM tasks (Wellman, Cross & Watson, 2001; Charman, Ruffman & Clements, 2002; Baron-Cohen, 2003; Wellman & Liu, 2004; Milligan, Astington & Dack, 2007).

The research did not found a significant correlation between the Eyes Test and somatic complaints. It seems that the ability to understand mental states from visual cues is independent from somatization.

The voice non verbal cues (e.g., tone, pauses, hesitations, rhythm) are more informative than visual information in social relationships (Wiseman, 1995). Children with frequent
somatic complaints, belonging to a normal population, can grasp mental states from external and observable cues (for example a child can recognize to be sad if he cries after he is smacked by his parent). Following theories of embodied cognition, supported from the discovery of the mirror neuron system (Gallese & Goldman, 1998; Wialliams et al., 2001; Gallese, 2003; Dapretto et al., 2006), children, with and without frequent somatic complaints, activates in their “body” the same emotion of the other when they observe the other’s emotional expression (Niedenthal, 2007).

The present study found that children with frequent somatic complaints failed more than children with few somatic complaints when they asked to recognize epistemic and emotional states from voices (as found in children with depression that misinterpret affective auditory stimuli; Emerson, Harrison & Everhart, 1999; Segrin, 2000).

The relationship of frequent somatic complaints and the Voice Test performance suggests that the impairment regarding the mindreading of their own emotions (that causes somatic complaints; Waller & Scheidt, 2006) is linked to the misinterpretation of others’ mental states (in Lucariello’s, 2004, words it shows the relationship between the intrapersonal and interpersonal ToM). As I summarized in the paragraph 1.4. ToM researchers are focalized most of all on others’ mental states understanding, but there is a lack of studies regarding the ability to impute mental states to oneself and the relationship that can be exist between one’s own and others’ mental states comprehension.

This result also suggest that the problem in children with psychosomatic symptoms is linked to the understanding not only of complex emotional states (e.g., Rieffe et al., 2007), but also of epistemic states and that both types of mental states can be implied in their emotion regulation difficulties.
3.4. Conclusion: Theory of Mind and emotion in children with frequent somatic complaints

Two studies were conducted in order to investigate the relationship between Theory of Mind and emotional difficulties in school age children.

The first research was aimed to study if ToM is linked to the risk to develop conduct disorder (externalizing problem) and depression, anxiety and somatization (internalizing problems). It found that children at risk of depression and somatization had poor performance in the Voice Test, an advanced ToM task based on vocal cues. These results are novel within the field of ToM studies for two reasons: they underline that it can be useful to investigate ToM not only in children with proclaimed pathologies but also in normal children at risk of psychological problems; moreover they point out for the first time the relationship between ToM and internalizing problems, that are characterized by emotional functioning difficulties, widening the ToM perspective from children with social impairments to those with emotional problems.

About depression, some ToM researchers showed that adult with this psychopathology are not able to make accurate judgments of others’ mental states (e.g., Inoue, Yamada & Kanba, 2006). The first study I presented in this chapter suggests that this difficulty can already exist during school age in children at risk to develop depression.

The link between mentalization and the risk of somatization has never been studied by ToM researchers. This novel relationship and the recent findings regarding emotion understanding difficulties in children with psychosomatic problems (Waller & Scheidt, 2006) suggested that a mentalization impairment could be found in children with frequent somatic complaints.
To investigate this subject a second research was conducted. It investigated specifically if the poor performance in ToM tasks is related to higher level of somatic complaints. It was found that children with frequent somatic complaints had difficulties to understanding complex epistemic and emotional mental states from voices. I think that this link could confirm that the psychosomatic disease testifies the “non-installation” of the mind in the body, that is the non-integration, within the body, of cognitions and emotions (Cerutti & Guidetti, 2007, p. 23).

This result marks out some aspects and suggest future research directions about ToM and emotional difficulties.

First, children with frequent somatic complaints not only cannot recognize emotions but also epistemic mental states, suggesting that probably their emotional problems are in deal with the understanding of a wide range of mental states and not only emotions. Following the recent idea, developed by Sharp, Croudace and Goodyer (2007), to study ToM not as present-absent, but as a mentalistic style, researches will be investigate if children with psychosomatic disease make specific mentalising bias about themselves and others.

Second, the problem in reading mental states in children with frequent somatic complaints seems to be not circumscribed to the self (intrapersonal process: e.g., misinterpretation of one’s own emotions), but also regards others’ mental states. The relation between intra- and interpersonal mental state understanding could be better investigate if ToM researchers move their attention also to the “reading” of one’s own mental states, creating instrument able to study also this ability.

Third, children with frequent somatic complaints can understand others’ mental states from visual cues, but find difficulty to read mental states from nonverbal cues of voices, that are not visible; this result will require a deeper investigation, in order to
verify, also through neuroscientific studies, how and why “look” differs to “listen” in children with and without somatization problems and in adults, too.

Moreover future researches should better explain the nature of relationship between ToM and somatic complaints: difficulty in ToM could be a prerequisite for the development of psychosomatic diseases? Or psychosomatic problems could affect the misinterpretation of mental states? Or, as I suggest, both can be considered part of a single coordinated system, so their relationship can take a variety of forms?

3.4.1. A new line of research: ToM and emotion regulation

The results of these two studies I think suggest also another new topic, that is not still analyzed in ToM studies: the relationship between ToM and emotion regulation. ToM is studied often in relation to social competence (e.g., Hughes & Leekam, 2004; Sharp, 2006), while its link with emotion regulation ability is less known. Recently Liebermann, Giesbrecht and Muller (2007) examined the contribution of ToM to individual differences in emotion regulation in preschool children. They found that, controlling for verbal ability, the relation between mentalization (false belief understanding) and the emotional control approached to significance: children who have a good ToM are also able to modulate their emotional responses.

As the Authors suggested an objective for future researches would be to clarify the relation between emotion regulation and ToM. This aim can be reach following two ways: the study of correlation between ToM measures and measures that require the child to use internal emotional control processes (Liebermann, Giesbrecht & Muller, 2007) and the study of ToM in children with emotional difficulties, like in the
researches presented in this chapter, where participants are normal children that can have some problems in emotion regulation.

The construct of emotion regulation has received a growing attention in these last years (Spinrad et al., 2006), but there has been little consensus regarding its definition (Thompson, 1994; Cole, Martin & Dennis, 2004); moreover, as Denham (1998; Denham et al., 2003) wrote, emotion regulation is inextricable from the experience, the expression and the understanding of emotions (see also Saarni, 1999; Miller et al., 2006). Because emotions are a dynamic organisation of physiological, cognitive and behavioral responses, occurring inside and outside awareness, the emotion regulation process involves all these three areas (Garber & Dodge, 1991; Walden & Smith, 1997). So the ability to understand one’s own emotions, to give meaning to observable indices of emotions (e.g., facial expressions, gestures, tone of voice) as well to the physiological components of the emotional arousal (e.g., heart rate), to select the appropriate behavior, are implied in emotion regulation (Walden & Smith, 1997). A wide definition of emotion regulation, even if it is still a “work in progress”, is that of Eisenberg and Spinrad (2004, p. 338): “emotion-related self-regulation as the process of initiating, avoiding, inhibiting, maintaining, or modulating the occurrence, form, intensity, or duration of internal feeling states, emotion-related physiological, attentional processes, motivational states, and/or the behavioral concomitants of emotion in the service of accomplishing affect-related biological or social adaptation or achieving individual goals”. Thus emotion regulation refers to changes associated with emotions, that are changes in the emotions themselves (Thompson, 1994) and in other psychological processes (e.g., social functioning, memory), and it is strictly interlaced with cognition and behavior (Cole, Martin & Dennis, 2004).
It is evident the role that the management of emotion has in children social competence (Hubbard & Coie, 1994; Eisenberg, 2002; Denham et al., 2003; Han & Kemple, 2006): those who are able to control impulses, delay gratification, reflect and monitor on one’s emotions have more likely the ability to be effective in the realization of social goals (so they are able, for example, to have friends, to be liked by others, to engage positive social interaction). It was found that children with difficulty in social adaptation are at risk for emotion regulation problems (e.g., Eisenberg et al., 1993; 1997, 2001, 2005; Rydell, Berlin & Bohlin, 2003; Spinrad et al., 2006).

Thus it can be suggested that ToM could be related both directly and indirectly (through emotion regulation) to children adaptive social functioning.

Many factors influence the development of emotion regulation, such as the child temperament and cognitive aspects (Diamond & Aspinwall, 2003; Morris et al., 2007). Two decades of empirical works found also that the attachment relationship, and more in general the family context (Morris et al., 2007), is related with the development of emotion regulation (Cassidy, 1994; Braumgart-Ricker et al., 2001; Mikulincer, Shaver & Pereg, 2003). In fact, as noticed van IJzendoorn (2007), one of the most important function of children’s attachment is to help to regulate overwhelming emotions, most of all negative emotions, therefore children can develop – within the mother-child relationship – adequate strategies to cope with these emotions: attachment is a scaffold for children’s emotion regulation.

It can be noted that also within ToM studies it was found that attachment relationship as an important role: it can affect mentalization development (e.g., Fonagy, Target, 1997).

Regarding specifically psychosomatic problems, researches found that children who suffer of somatic symptoms not explained by a medical diagnosis, have difficulty in
emotion regulation, in particular they have poor emotion awareness and often fail to identify emotions and to differentiate them from physiological states (e.g., Jellesma et al., 2006; Rieffe et al., 2007).

They are not able to consider their own emotions as mental states, because focus their attention on the bodily sensations of the emotional arousal, misinterpreting physical signals and considering them as somatic symptoms, without links with the emotional mental state (e.g., Waller & Scheidt, 2006).

\[16\]

In particular, there are different theoretical models that link cognitive deficits in the processing of emotions to somatization (Waller & Scheidt, 2006). One hypothesis was developed in the field of alexithymia research (reviews: Taylor & Bagby, 2004; Toni, 2007). Alexithymia is considered to be a personality trait, it is defined by: difficulty to identify feelings and distinguish between feelings and the bodily sensations of emotional arousal; difficulty to describe feelings to other people; constricted imaginative processes and paucity of fantasies; a stimulus-bound, externally oriented cognitive style (Nemiah & Sifneos, 1970; Nemiah, Freyberger & Sifneos, 1976; Taylor & Taylor, 1997). It is conceptualized as an emotion regulation disorder reflecting deficits in the cognitive and interpersonal regulation of emotion (Taylor, Bagby & Parker, 1997; about the link between somatization and alexithymia see also: Cohen, Auld & Brooker, 1994; Cox et al., 1994; Lundh & Simonsson-Sarneki, 2001; De Gucht & Heiser, 2003). The alexithymic persons’ limited capacities to regulate emotions through cognitive processes (e.g., identify and understanding emotions, recognize their causes) lead them to focus on bodily sensations, that accompany the emotional arousal, so they amplify or misinterpret these physical signals (Taylor, Bagby & Parker, 1997), not linking them to emotions.

Another theoretical model underlines the lack of emotional awareness in somatization disorders (Lane & Schwartz, 1987; Lane et al., 1990). Lane and Schwartz (1987) proposed a cognitive-developmental model of emotional awareness and described five levels of emotional awareness: awareness of bodily sensations, of the action tendencies connected to emotions, of an emotion, of blends of emotion and, the more complex cognitive schemata, the awareness of combinations of blends of emotion. The cognitive schemata for processing emotions are not developed in persons suffering of psychosomatic symptoms, that are not able to elaborate bodily emotional arousal in emotions consciously felt, but are only aware of physiological cues. Finally, somatization can be considered as the result of dissociation between the sub-symbolic and symbolic components of the emotional schemata (Bucci, 1997). Bucci (1997), in her multiple code theory, focalized on emotions representation in non-verbal and verbal channels. Non-verbal ones include both sub-symbolic (sensory, somatic, visceral and motor modalities) and symbolic processes (imagery). The emotion somatic and motor arousals are isolated from the cognitive activation, so persons cannot regulate emotions and experience a prolonged physical activation, that can be associated with the development of psychosomatic problems.
Within ToM perspective, Fonagy and colleagues (2002) hypothesized that affect regulation\(^{17}\), that is present from birth within the mother-child dyadic system (e.g., Lavelli, 2007), becomes more mature and sophisticated thanks to mentalization development (see also paragraph 1.2.2.). Following them, it can be argued that the relationship between emotion regulation ability and ToM is complex, they influence each other. So somatic complaints, that testify a poor emotion regulation ability, could be interlaced with mentalization in a complex way and probably they have a common origin in the child-caregivers relationship (Fonagy et al., 2002; Morris et al., 2007).

It could be interesting to study how ToM, somatic complaints, emotion regulation and social competences are related. Till now researches found that: ToM is important to social development but it is probably not sufficient to have positive social behavior (social adaptation) (Astington, 2003); children with somatic complaints have problems both in emotion regulation and in social life (Garralda, 1999); ToM seem to be related to somatic complaints (these researches); and ToM is probably linked to emotion regulation (both from a psychoanalytic view, Fonagy et al., 2002; and a cognitive view: Liotti, 2001; Colle, 2007).

In conclusion, in this chapter I present a novel and I think fruitful line of research within the field of study of ToM, regarding the link between social understanding and emotional functioning in children with emotional difficulties. Certainly it opens more questions than those it answered, in fact starting from the hypothesis to investigate ToM in internalizing problems, the chapter ends suggesting a complex view of ToM and emotional functioning in the development of children individual differences.

\(^{17}\) Affect regulation is the emotion regulation in the child-caregiver attachment relationship (Fonagy et al., 2002; see also chapter 1).
Conclusion

This PhD work had the aim to widen Theory of Mind developmental perspective on new topics, trying to pay attention to the theoretical (e.g., ToM definition) and methodological (e.g., type of instruments and methodological choice) questions that this open view suggests.

Trying to schematize the novelties introduced by the researches on ToM here presented, they are:

- focalization on Italian typical school age children;
- creation of a new advanced ToM task based on vocal cues;
- focalization on children’s psychological risk;
- attention to emotional difficulties, typical of internalizing problems.

It can be said, at the end of the thesis, that these new lines of researches suggest other new areas of investigations, that reflect a complex view of ToM, such as the relationship between:

- one’s own and others’ mental state understanding;
- the understanding of epistemic and emotional mental states from visual cues and from vocal cues, but also from narrative stimuli;
- ToM and emotional functioning, in particular emotional regulation;

and also the role of ToM in children socio-emotional difficulties and the possibility to use ToM in order to identify precociously children’s psychological problems.

In the first chapter I reviewed some important topics, that are still open issues within the ToM field: the various theoretical approaches and the ToM definitions they
adopted; the link between mentalization and social/emotional functioning; the
development of ToM and its assessment. The researches I presented here were in deal
with all these aspects, suggesting the importance to adopt a wide definition of ToM,
understood and assessed as epistemic and emotional states comprehension and linked
to the social and emotional functioning.
In particular, the second chapter had the aim to provide a new advanced ToM task for
typical Italian school age children, based on vocal cues: the *Voice Test*. It was described
its creation, validation and standardization, that involved a total of 983 children,
showing its good psychometric qualities (internal and test-retest reliability and construct
validity) and giving normative data for children aged from 6,5 to 11,4 years. It provided
to consider auditory stimuli in mental states understanding and to analyze deeper the
increasing mentalization ability in school age children.
The third chapter presented two studies that had the general aim to analyze the
relationship between ToM and emotional difficulties in a normal school age children
population. The first study, in which participated 112 children, found that children at
risk to develop depression and somatization had difficult to understand mental states
from vocal cues, while children at risk to develop conduct disorders and anxiety did not
show mentalization impairment.
The second study investigated deeper the link between ToM, assessed using both
classical and advanced tasks, and psychosomatic disease in 96 children, founding that
children with frequent somatic complaints had difficult to understand complex mental
states from voices. In conclusion they confirmed the idea, supported also by studies on
emotion understanding conducted out of the field of ToM, that mentalization can deal
also with internalizing problems and the emotional functioning.
References


ALBANESE O., MOLINA P. (Eds.) (in press), *Lo sviluppo della comprensione delle emozioni e la sua valutazione*, Unicopli, Milano.


BATTISTELLI P. (1992), La rappresentazione della soggettività, Franco Angeli, Milano.


Borg E. (2007), *If mirror neurons are the answer, what was the question?,* in “Journal of Consciousness Studies”, 14, pp. 5-19.


Call J. (2007), *Apes know that hidden objects can affect the orientation of other objects*, in “Cognition”, 105, pp. 1-25.


CHARMAN T., CARROLL F., STURGE C. (2001), Theory of mind, executive function, and social competence in boys with ADHD, in “Emotional and Behavioral Difficulties”, 6, pp. 31-49.


CORICELLI G. (2005), Two-levels of mental states attribution: from automaticity to voluntariness, in “Neuropsychologia”, 43, pp. 294-300.


of children's social functioning from regulation and emotionality, in “Child Development”, 68, pp. 642–64.


HARRIS P. L. (1992), From simulation to folk psychology: The case for development, in “Mind and Language”, 7, pp. 120-44.


Legerstee M., Barillas Y. (2003), *Sharing attention and pointing to objects at 12 months: is the intentional stance implied?,* in “Cognitive Development”, 18, pp. 91-110.


Liszowski U. (2005), *Human twelve-month-olds point cooperatively to share interest with and provide information for a communicative partner*, in “Gesture”, 5, pp. 135-54.


LIVERTA SEMPIO O., MARCHETTI A. (2001b), Il compito di falsa credenza di primo ordine: lo spostamento inatteso, Theory of Mind Research Unit, Department of Psychology, Catholic University of the Sacred Heart, Milan.

LIVERTA SEMPIO O., MARCHETTI A. (2001c), Il compito di falsa credenza di primo ordine: la scatola ingannevole, Theory of Mind Research Unit, Department of Psychology, Catholic University of the Sacred Heart, Milan.


LIVERTA SEMPIO O., MARCHETTI A., CASTELLI I. (2003), Test degli occhi – versione bambini, Theory of Mind Research Unit, Department of Psychology, Catholic University of the Sacred Heart, Milan.


LIVERTA SEMPIO O., MARCHETTI A., LECCISO F. (Eds.) (2005a), Teoria della mente tra normalità e patologia, Raffaello Cortina, Milano.

LIVERTA SEMPIO O., MARCHETTI A., LECCISO F. (2005b), Faux Pas: traduzione italiana, Theory of Mind Research Unit, Department of Psychology, Catholic University of the Sacred Heart, Milan.

LIVERTA SEMPIO O., MARCHETTI A., CASTELLI I., LECCISO F., PEZZOTTA C. (2005), Mentalizzazione e competenza sociale. La comprensione della falsa credenza nello sviluppo normale e patologico, Franco Angeli, Milano.


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, in “Child
Development”, 78, pp. 622-46.

1-14.

MORRIS S., SILK J., STEINBERG L., MYERS S., ROBINSON L. (2007), The Role of the Family
Context in the Development of Emotion Regulation, in “Social Development”, 16, pp. 361-
88.

Development”, 72, pp. 834-43.

1-14.

MORRIS S., SILK J., STEINBERG L., MYERS S., ROBINSON L. (2007), The Role of the Family
Context in the Development of Emotion Regulation, in “Social Development”, 16, pp. 361-
88.

Development”, 72, pp. 834-43.

MOSES L., BALDWIN D., ROSICKY J., TIDBALL G. (2001), Evidence for referential
understanding in the emotions domain at 12 and 18 months, in “Child Development”, 72,
pp. 718-35.

development, and children’s interpretation of anxiety-related physical symptoms, in “Behaviour
Research and Therapy”, 45, pp. 2121-32.

MURIS P., STEERNEMAN P., MEESTERS C., MERCKELBACH H., HORSELENBERG R., VAN
DEN HOGEN T., VAN DONGEN L. (1999), The TOM Test: A new instrument for assessing
theory of mind in normal children and children with pervasive developmentl disorders, in

NADEL J., MUIR D. (Eds.) (2005), Emotional Development: Current and future research
directions, Oxford University Press, Oxford.

NAGY E. (2006), From imitation to conversation: The first dialogues with human neonates, in

NELSON J., STAGE S., DUPPONG-HURLEY K., SYNHORT L., EPSTEIND M. (2007), Risk
Factors Predictive of the Problem Behavior of Children At Risk for Emotional and Behavioral
Disorders, in “Exceptional Children”, 73, pp. 367-79.

In O. W. Hill (Ed.), Modern trends in psychosomatic medicine, vol. 2, Butterworths,
London, pp. 26-34.

NEMIAH J. C., FREYBERGER H., SIFNEOS P. E. (1976), ALEXITHYMIA: A view of the
psychosomatic process. In O.W. Hill (Ed.), Modern Trends in Psychosomatic Medicine, vol. 3,

NICHOLS S., STICH S. (2003), Mindreading: an integrated account of pretence, self-awareness, and
understanding of other minds, Oxford University Press, New York.


statements and hold beliefs, in “Educational Psychologist”, 28, pp. 7-23. Italina
translation in O. Liverta Sempio, A. Marchetti (Eds.) (2005), Il pensiero dell’altro,
Raffaello Cortina, Milano, 1995, pp. 429-446.


Peterson C., Slaughter V. (2003), Opening windows into the mind: mothers’ preferences for mental state explanations and children’s theory of mind, in “Cognitive Development”, 18, pp. 399-429.


PRATT C., BRYANT P. (1990), Young children understand that looking leads to knowing (so long as they are looking into a single barrel), in “Child Development”, 61, pp. 973-983.


STEELE H., STEELE M., CROFT C., FONAGY P. (1999), Infant-Mother Attachment at One Year Predicts Children’s Understanding of Mixed Emotions at Six Years, in “Social Development”, 8, pp. 161-78.


STERN D. N. (1998), Le interazioni madre-bambino nello sviluppo e nella clinica, Raffaello Cortina, Milano.


STRIANO T., BERTIN E. (2005), Social-cognitive skills between 5 and 10 months of age, in “British Journal of Developmental Psychology”, 23, pp. 559-68.


SUTTON J. (2001), Bullies: Thugs or Thinkers?, in “The Psychologist”, 14, pp. 530-34.


Appendix

THE VOICE TEST

Materials
- The CD of the Voice Test, containing 21 spoken phrases, and a CD player
- The answer sheet of the Voice Test and the gender recognition task
- The Glossary

Mental states associated to spoken phrases

0) (Trial item) Spaventata → “Tu, tu sospetti qualcuno?”
1) Preoccupata → “Prego, dobbiamo andare”
2) Sta pensando a qualcosa → “Ho visto il tuo progetto”
3) Dubbiosa → “Ma cosa intendi?”
4) Seria → “No davvero”
5) Sta pensando a qualcosa → “Devo spiegarti che cosa intendevo”
6) Dubbiosa → “Non ho capito bene”
7) Che ha deciso qualcosa → “Parlerò con tuo fratello”
8) Un po’ preoccupata → “Era chiaro che ci saremmo dimenticati”
9) Interessata → “Non mi avevi detto di avere una sorella”
10) Scontenta → “Ma io avevo sperato”
11) Interessata → “Veramente se n’è andato?”
12) Sta pensando a qualcosa → “Ormai potremmo decidere”
13) Sicura di qualcosa → “C’è una cosa che Luigi vuol dire”
14) Nervosa → “C’è...c’è qualcosa che voglio chiederti”
15) Dubbiosa → “E’ stato assodato”
16) Supplichevole → “Giuro che ce l’ho”
17) Dispiaciuta → “Sembra che non sia arrivato in tempo”
18) Seccata → “E tientela”
19) Seria → “Penso piuttosto che abbiamo alcune cose di cui discutere”
20) Confusa → “Ma che cosa intendi?”

*Instructions (in Italian)*

Su questo registratore sono incise le voci di diverse persone che pronunciano delle frasi. Ti/vi farò ascoltare una voce alla volta. Tu devi/voi dovete ascoltarla attentamente poi scegliere su questo foglio, nella riga corrispondente, la parola che meglio secondo te/voi esprime che cosa sta pensando o provando la persona che ha appena parlato.

Prima di ascoltare la voce leggerò le 4 parole tra cui dovrai/dovrete scegliere cosa pensa o cosa prova la persona che parlarà. Se non sei sicuro/non siete sicuri del significato di qualche parola, possiamo guardare insieme sul Glossario cosa significa.

Devi/dovete fare attenzione perché ti/vi farò ascoltare solo una volta la voce, poi avrai/avrete tutto il tempo che vuoi/volete per scegliere la risposta. Quando hai/avete fatto, passerò alla voce successiva.
Facciamo una prova insieme. Ascolta questa voce. Secondo te/voi, come si sente o cosa ha in mente la persona che ha parlato? E’ gelosa, spaventata, incurante o prova odio? (ascolto dell’item di prova).

Alcune volte ti/vi sembrerà facile rispondere, altre volte potrebbe sembrare difficile, quindi non ti preoccupare/vi preoccupate se non sarà sempre facile scegliere la parola migliore. Leggerò io tutte le parole tra cui devi/dovete scegliere.

(Al termine della prova, viene introdotto il test di controllo) Ora ti/vi farò riascoltare le stesse voci che hai appena sentito, stavolta ti chiedo di indicare se secondo te/voi è una voce di un maschio o di una femmina, mettendo una crocetta sulla lettera M se pensi sia un maschio e sulla F se pensi sia una femmina.

Facciamo una prova insieme. Ascolta questa voce. Secondo te è una voce di maschio o di femmina? (ascolto dell’item di prova)

Fai/fate molta attenzione, perché ora non fermerò più il cd dopo ogni frase, dovrai/dovrete mettere la crocetta su M o F subito dopo aver sentito la voce, perché poi sentirai/sentirete subito un’altra voce.
### Answer sheet of the Voice Test

<table>
<thead>
<tr>
<th></th>
<th>GELOSA</th>
<th>SPAVENTATA</th>
<th>INCURANTE</th>
<th>STA PROVANDO ODIO</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>STA PROVANDO ODIO</td>
<td>SGARBATA</td>
<td>PREOCCUPATA</td>
<td>TEDIATA</td>
</tr>
<tr>
<td>2</td>
<td>SECCATA</td>
<td>STA PROVANDO ODIO</td>
<td>SORPRESA</td>
<td>STA PENSANDO A QUALCOSA</td>
</tr>
<tr>
<td>3</td>
<td>GENTILE</td>
<td>TIMIDA</td>
<td>DUBBIOSA</td>
<td>TRISTE</td>
</tr>
<tr>
<td>4</td>
<td>CONFUSA</td>
<td>SCHERZOSA</td>
<td>TRISTE</td>
<td>SERIA</td>
</tr>
<tr>
<td>5</td>
<td>STA PENSANDO A QUALCOSA</td>
<td>SCONVOLTA</td>
<td>ECCITATA</td>
<td>FELICE</td>
</tr>
<tr>
<td>6</td>
<td>DUBBIOSA</td>
<td>AMICHEVOLE</td>
<td>SCHERZOSA</td>
<td>INCURANTE</td>
</tr>
<tr>
<td>7</td>
<td>CHE HA DECISO QUALCOSA</td>
<td>SCHERZOSA</td>
<td>SORPRESA</td>
<td>TEDIATA</td>
</tr>
<tr>
<td>8</td>
<td>ARRABBIATA</td>
<td>AMICHEVOLE</td>
<td>SGARBATA</td>
<td>UN PO' PREOCCUPATA</td>
</tr>
<tr>
<td>9</td>
<td>ARRABBIATA</td>
<td>SOPRAPENSIERO</td>
<td>TRISTE</td>
<td>INTERESSATA</td>
</tr>
<tr>
<td></td>
<td>GENTILE</td>
<td>SORPRESA</td>
<td>SCONTENTA</td>
<td>ECCITATA</td>
</tr>
<tr>
<td>---</td>
<td>-------------</td>
<td>----------</td>
<td>-----------</td>
<td>-----------</td>
</tr>
<tr>
<td>10</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>11</td>
<td>INTERESSATA</td>
<td>SCHERZOSA</td>
<td>INCURANTE</td>
<td>FELICE</td>
</tr>
<tr>
<td>12</td>
<td>SCHERZOSA</td>
<td>GENTILE</td>
<td>SORPRESA</td>
<td>STA PENSANDO A QUALCOSA</td>
</tr>
<tr>
<td>13</td>
<td>SORPRESA</td>
<td>SICURA DI QUALCOSA</td>
<td>SCHERZOSA</td>
<td>FELICE</td>
</tr>
<tr>
<td>14</td>
<td>SCHERZOSA</td>
<td>INCURANTE</td>
<td>NERVOSA</td>
<td>DISPIACIUTA</td>
</tr>
<tr>
<td>15</td>
<td>SI VERGOGNA</td>
<td>ECCITATA</td>
<td>DUBBIOSA</td>
<td>APPAGATA</td>
</tr>
<tr>
<td>16</td>
<td>ORRIPILATA</td>
<td>SUPPLICHEVOLE</td>
<td>SI STA SCUSANDO</td>
<td>GENTILE</td>
</tr>
<tr>
<td>17</td>
<td>DISPIACIUTA</td>
<td>SCONVOLTA</td>
<td>INCURANTE</td>
<td>SERIA</td>
</tr>
<tr>
<td>18</td>
<td>CONFUSA</td>
<td>SORPRESA</td>
<td>SECCATA</td>
<td>FELICE</td>
</tr>
<tr>
<td>19</td>
<td>SI VERGOGNA</td>
<td>CONFUSA</td>
<td>AMICHEVOLE</td>
<td>SERIA</td>
</tr>
<tr>
<td>20</td>
<td>CONFUSA</td>
<td>ECCITATA</td>
<td>TEDIATA</td>
<td>PREPOTENTE</td>
</tr>
</tbody>
</table>
### Answer sheet of the gender recognition task

<p>| | | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>M</td>
<td>F</td>
</tr>
<tr>
<td>1</td>
<td>M</td>
<td>F</td>
</tr>
<tr>
<td>2</td>
<td>M</td>
<td>F</td>
</tr>
<tr>
<td>3</td>
<td>M</td>
<td>F</td>
</tr>
<tr>
<td>4</td>
<td>M</td>
<td>F</td>
</tr>
<tr>
<td>5</td>
<td>M</td>
<td>F</td>
</tr>
<tr>
<td>6</td>
<td>M</td>
<td>F</td>
</tr>
<tr>
<td>7</td>
<td>M</td>
<td>F</td>
</tr>
<tr>
<td>8</td>
<td>M</td>
<td>F</td>
</tr>
<tr>
<td>9</td>
<td>M</td>
<td>F</td>
</tr>
<tr>
<td>10</td>
<td>M</td>
<td>F</td>
</tr>
<tr>
<td>11</td>
<td>M</td>
<td>F</td>
</tr>
<tr>
<td>12</td>
<td>M</td>
<td>F</td>
</tr>
<tr>
<td>13</td>
<td>M</td>
<td>F</td>
</tr>
<tr>
<td>14</td>
<td>M</td>
<td>F</td>
</tr>
<tr>
<td>15</td>
<td>M</td>
<td>F</td>
</tr>
<tr>
<td>16</td>
<td>M</td>
<td>F</td>
</tr>
<tr>
<td>17</td>
<td>M</td>
<td>F</td>
</tr>
<tr>
<td>18</td>
<td>M</td>
<td>F</td>
</tr>
<tr>
<td>19</td>
<td>M</td>
<td>F</td>
</tr>
<tr>
<td>20</td>
<td>M</td>
<td>F</td>
</tr>
</tbody>
</table>
Glossary

**AMICHEVOLE**: affabile, confidenziale, amicale

Esempio: La nuova vicina di casa di Maria è una persona davvero amichevole perché quando incontra Maria le parla sempre volentieri e amabilmente.

**APPAGATA**: si sente soddisfatta

Esempio: Cristina aveva desiderato così tanto di visitare Roma che quando finalmente ci è andata si è sentita davvero appagata.

**ARRABBIATA**: irata, rabbiosa, adirata

Esempio: La mamma è arrabbiata perché Giovanni non ha messo in ordine la sua cameretta.

**CHE HA DECISO QUALCOSA**: ha preso una decisione, risoluta, determinata

Esempio: Dopo aver riflettuto a lungo, Massimo ha deciso che non parteciperà al torneo di tennis.

**CONFUSA**: disorientata, incerta, perplessa

Esempio: Lisa era stata interrotta così tante volte mentre parlava che alla fine si è sentita confusa e ha taciuto.

**DISPIACIUTA**: addolorata, rammaricata, rincresciuta

Esempio: Francesco è dispiaciuto perché il suo amico Sandro si è fatto male giocando a pallone.

**DUBBIOSA**: fatica a credere a qualcuno o in qualcosa, scettica

Esempio: Marco giura di non aver rotto il vaso, ma la zia è dubbia e fatica a credergli.

**ECCITATA**: esaltata, animata, vivace

Esempio: E' la mattina di Natale e Luca è eccitato nell'aprire i suoi regali.

**FELICE**: beata, allegra, contenta
Esempio: Elisa è **felice** perché ha preso un bel voto nel compito in classe.

**GELOSA**: ingelosita, sospettosa, timorosa, possessiva

Esempio: Sabrina era **gelosa** che Anna, la sua amica del cuore, giocasse con le altre compagne.

**GENTILE**: affabile, amabile, beneducata

Esempio: Giulia è molto **gentile** quando qualcuno le chiede informazioni: risponde con precisione e sorride.

**INCURANTE**: tranquilla, non si fa troppi problemi

Esempio: Paola e Lucia discutevano sempre più animatamente, ma Carla incurante le lasciava fare.

**INTERESSATA**: vuole saperne di più, curiosa

Esempio: Dopo aver visto un film sugli orsi, Claudia è molto **interessata** a questi animali: vuole sapere tutto su di loro.

**NERVOSA**: irritabile, tesa, suscettibile

Esempio: Massimiliano quel mattino era **nervoso** perché doveva essere interrogato.

**ORRIPILATA**: terrorizzata, terrificata

Esempio: Maria ha molta paura e schifo dei ragni: quando ne vede uno è orripilata.

**PREOCCUPATA**: ansiosa, impensierita, apprensiva

Esempio: Quando il suo gattino non è più tornato a casa, Maria era preoccupata.

**PREPOTENTE**: arrogante, insolente, impone a ogni costo la propria volontà

Esempio: Alessandro è **prepote** con il suo fratellino: gli ruba sempre tutti i suoi giocattoli.
SCHERZOSA: giocherellona, divertente, spiritosa
   Esempio: Gabriele era sempre scherzoso con i suoi amici: li faceva divertire piacevolmente.

SCONTENTA: infelice, insoddisfatta, inappagata
   Esempio: In mensa non c’è mai nulla che gli piaccia, quindi Luciano è sempre scontento del menù.

SCONVOLTA: agitata, scombussolata, turbata
   Esempio: Mara è sconvolta perché ha appena saputo che suo fratello ha avuto un incidente.

SECCATA: stufa, infastidita, irritata
   Esempio: Il papà è seccato quando qualcuno lo disturba mentre sta leggendo il suo giornale preferito.

SERIA: austera, solenne, grave
   Esempio: Il preside era molto serio quando rimproverò gli alunni per aver rovinato i banchi.

SGARBATA: maleducata, scortese, villana
   Esempio: Maria è sgarbata con chi la disturba mentre studia.

SI STA SCUSANDO: chiedere perdono, discolparsi
   Esempio: Luisa ha rotto un vaso di cristallo che piaceva alla mamma ed ora si sta scusando con lei dicendo che non l’ha fatto apposta.

SI VERGOGNA: prova vergogna, si sente in imbarazzo, si sente avvilita
   Esempio: La signora Anna si vergognò quando si accorse di essere uscita di casa con i bigodini ancora in testa.

SICURA DI QUALCOSA: certa, non ha dubbi
Esempio: Alessandra risponde a una domanda difficile di geografia sicura che la sua risposta sia giusta.

**SOPRAPPENSIERO:** sta pensando ad altro, è immersa nei suoi pensieri

Esempio: Mentre usciva di casa Clara era talmente soprappensiero che si è dimenticata di chiudere la porta a chiave.

**SORPRESA:** stupita, sconcertata

Esempio: Giovanna è sorpresa per un regalo inaspettato.

**SPOVENTATA:** intimorita, ha paura di qualche cosa o di qualcuno, è presa da terrore

Esempio: Laura è spaventata dai tuoni e dai lampi del temporale.

**SPERANZOSA:** piena di speranza, fiduciosa, aspetta di ottenere ciò che desidera

Esempio: Stefano adora il circo ed è speranzoso che il papà torni in tempo per poterci andare.

**STA PENSANDO A QUALCOSA:** riflettere su qualche cosa, meditare su qualcosa, formare nella mente dei pensieri su qualche cosa

Esempio: Alice sta pensando a come organizzare la sua festa di compleanno.

**STA PROVANDO ODIO:** detesta qualcuno, ce l’ha molto con qualcuno

Esempio: Giorgio sta provando odio per il compagno che ha fatto la spia.

**SUPPLICHEVOLE:** che implora, che prega per avere qualcosa

Giuseppe ama molto i gattini e da qualche giorno chiede supplichevole alla mamma di averne uno in regalo.

**TEDIATA:** prova noia, fastidio, si sta stufando

Esempio: Il film che Alessia sta guardando non è per niente divertente e quindi lei è proprio tediata.

**TIMIDA:** chiusa, introvera, vergognosa
Esempio: Caterina, che è **timida**, diventa subito rossa quando qualcuno le chiede qualcosa.

**TRISTE**: abbattuta, afflitta, malinconica

Esempio: Mario è **triste** perché il suo compagno di banco ha cambiato scuola.