Journal of Autism and Developmental Disorders (2019) 49:3494–3503
https://doi.org/10.1007/s10803-019-04064-3

ORIGINAL PAPER

Theory of Mind Development in Children with Visual Impairment: The Contribution of the Adapted Comprehensive Test ToM Storybooks

Gloriana Bartoli[1,2,3] · **Daniela Bulgarelli**[4] · **Paola Molina**[4]

Published online: 22 May 2019

Abstract

Research that focused on Theory of Mind (ToM) development in blind children showed that they were delayed, but not permanently deficient, in various types of false belief tasks. More recent studies reported first evidence of typical ToM development in blind children and suggested that more comprehensive tools to evaluate ToM had to be used. The current paper analyzed ToM development in blind children, using the adapted version of the ToM Storybooks; this is a standardized comprehensive test developed to provide a reliable and stable measurement, in comparison with the false belief tasks. Results showed that blind children's ToM performances were very similar to the ones of matched typically developing children, matched on chronological age and gender. The current finding supported the importance of the use of a more comprehensive tool to assess ToM in atypical population.

Keywords Atypical development · Blindness · ToM components · First-order false belief task · Adapted test

Theory of Mind (ToM) is part of the socio-cognitive competence and refers to the ability to understand, predict and interpret one's own and others' behavior, thanks to the attribution of mental states such as desires, beliefs and thoughts to people (Mitchell 1997). The most widespread way to assess ToM consists in first-order false belief (FB) tasks. In the original version by Wimmer and Perner (Maxi task; 1983) and in the following revised version by Baron-Cohen et al. (Sally-Anne task; 1985), the child is told a story where an unexpected change happens: the protagonist puts an object in a box and then leaves the room; a second character moves the object into another box; the protagonist comes back and the child is asked where s/he will search for the object and where s/he thinks the object is. Typically, 3-year-olds systematically fail the task whereas, from 4 years of age, the probability of providing the correct answer goes from below chance to above chance. Several variants of the FB task have been proposed, that could be categorized into three types: location or unexpected change task (i.e. Sally-Anne task; Baron-Cohen et al. 1985), content or unexpected content task (i.e. Smarties task, in which a box does not contain what it is meant to contain; Perner et al. 1987) and identity or misleading appearance task (i.e. a sponge looking like a rock, a task in which an object with a deceptive identity is presented; Flavell 2004). Moreover, different versions of the same task were used, involving for instance real people or puppets, requiring verbal or non-verbal answers from the child. A meta-analysis showed that the developmental trend of typically developing (TD) children's responses does not depend on any of these variants, such as the type of task, type of questions, nature of the protagonist or the object (Wellman et al. 2001, 2018). Second-order FB task was introduced to examine people's belief about others' belief (i.e. "John thinks that Mary thinks that…", Perner and Wimmer 1985). Children passing the first-order FB task failed to pass the second-order FB task up to 7–9 years of age, thus the latter was used as a measure of more complex

✉ Daniela Bulgarelli
daniela.bulgarelli@unito.it

Gloriana Bartoli
gloriana.bartoli@gmail.com

Paola Molina
paola.molina@unito.it

1 Dipartimento di Psicologia, Università degli Studi di Pavia, Pavia, Italy

2 Fondazione Robert Hollman, Cannero Riviera, VB, Italy

3 Present Address: CAMHS: Child and Adolescent Mental Health Service, ADHB, Auckland District Health Board, and Private Practice in Auckland, Auckland, New Zealand

4 Università degli Studi di Torino, Dipartimento di Psicologia, Turin, Italy

ToM ability; but memory and executive functioning abilities seemed to mainly explain children's performances in the second-order FB task (Sullivan et al. 1994).

ToM development has been investigated in both typically developing children and in children with different types of disabilities or developmental disorders, such as autism spectrum disorder, specific language impairment, deafness and blindness, in order to identify which processes contribute to its development across life and what could account for differences among individuals. ToM development of deaf children born in oral families is delayed compared to that of typically developing children and deaf children born in signing families: in the first case, infants are not exposed to sign language for several months since their birth and the delay in being included in conversational exchanges negatively impacts on the development of both communication and ToM (Peterson 2009). The literature reported that children with autism, who show atypical interaction and communication skills, also show a delay in ToM compared to TD children when first-order and second-order FB tasks were used (Baron-Cohen et al. 1985; Yirmiya et al. 1998). Recently, a cross-sequential study involved 3- to 13-year old TD children, children with autism or deafness who were administered at Time 1 and at Time 2 (18 months later) three first order FB tasks and the six-step ToM Scale (including tasks on diverse desires, diverse beliefs, knowledge access, false belief, hidden emotion, sarcasm; derived from the ToM Tasks by Wellman et al. 2011). Although deaf and autistic children obtained lower performances than their TD peers both at Time 1 and 2, the three groups showed a similar progress in ToM competence and individual differences among the children remained stable (Peterson and Wellman 2018). Another recent study reported that high-functioning autistic children and adolescents between 6 and 20 years performed equally to TD peers, but they still might show limited ToM abilities in everyday social interactions (Scheeren et al. 2013). Other research reported that the delay in ToM development of children with specific language impairment is associated with the severity of the disorder: children with phonological disorders (i.e. difficulties in properly forming the sounds of words) showed performances similar to those of typically developing children, whereas children with receptive-expressive disorders (i.e. difficulties in understanding and producing words and sentences) showed performances similar to those of autistic children (Bulgarelli and Molina 2013; Shields et al. 1996). Thus, delays in language and communication seem to be associated with delays in ToM.

The current paper specifically focuses on visually-impaired (VI) children, who are reported to manifest ToM delay and autistic-like behaviors (Fraiberg 1977; Bartoli et al. 2010; for a review Williams et al. 2014). Some studies compared "special conditions closely resembling autism", such as these of children with VI, to test whether "primary deficit or computational mechanisms are responsible for the clinical picture" (Hobson 2004, p. 190). The absence of visual co-orientation starting from early social and communicative interactions could explain ToM delay in blind children (Hobson 1990; Minter et al. 1998; Green et al. 2004). Roch-Levecq (2006) argued that blind children lacked a primitive form of relatedness to others, because they were not involved in a mutual feedback loop of affect-mirroring that relied on the ability to see, described by Gergely and Watson (1996); this would negatively affect ToM development. This kind of deprivation could possibly explain the reason why autistic-like behaviors are observed in blind children (Pérez-Pereira and Conti-Ramsden 2005).

Since 1995, few papers focused on ToM development in VI children. Most of the studies showed that children with congenital visual impairment were delayed of 4–7 years, but not permanently deficient, in various types of first order FB tasks, with the same percentage of blind children passing the FB task up to the age of 12 instead of four within the group of sighted children (Begeer et al. 2014; Brambring and Asbrock 2010; Green et al. 2004; McAlpine and Moore 1995; Minter et al. 1998; Peterson et al. 2000; Roch-Levecq 2006); this paired with the delay in perspective taking achieved around the age of 12 in the visual impaired sample (Farrenkopf and Davidson 1992). Nevertheless, recently, first evidence of typical ToM development in blind children was reported. Pijnacker et al. (2012) administered to blind children's several auditory tasks: two first-order FB tasks (i.e. unexpected contents), two second-order FB stories, and five stories tacking advanced ToM (i.e. lie, white lie, figure of speech, joke and irony). The research showed that the visually impaired children's performances did not differ from those of the control group children, matched on gender, age and verbal IQ. This result was coherent with Bedny et al. (2009) who reported that blind adults used the same neural network for ToM processing as sighted people; the localization and the selection of the network components were also similar. These recent results suggested that tools and measures including different tasks from the first-order FB are needed to evaluate ToM in blind children.

The studies on blind children's ToM deepened the role of some factors that could explain inter-individual differences in ToM performances. The types of FB tasks or the manipulations of the FB tasks were correlated with children's performance in initial studies (Minter et al. 1998), but this result was not confirmed in following studies (Brambing and Asbrock 2010; Green et al. 2004; Peterson et al. 2000). Brambring and Asbrock (2010) and Peterson et al. (2000) reported that differences were not associated to gender. Peterson et al. (2000) found no differences due to physical or learning disabilities. Green et al. (2004) excluded an association between type of school attended by the child (special vs mainstream) and FB performances. Nevertheless, an agreement was not reached about some other factors that could account for inter-individual differences in blind children's ToM performances. Verbal IQ and verbal mental age measured through the Wechsler's scales were positively associated with unexpected content and unexpected transfer FB tasks in Green et al. (2004), whereas verbal ability measured by syntactic complexity did not correlate with performances in the unexpected content tasks in Roch-Levecq (2006). Chronological age was positively associated with first-order FB tasks performance in Brambing and Asbrock (2010), Peterson et al. (2000), Pijnacker et al. (2012) and Roch-Levecq (2006), but not in Green et al. (2004). An association between ToM scores and visual acuity was retrieved in Roch-Levecq (2006) but not in Pijnacker et al. (2012). Finally, the role of type of visual diagnosis or degree of severity on ToM development was not clear: Green et al. (2004), Minter et al. (1998) and Peterson et al. (2000) argued that an association between these factors and FB performances is not present. Yet, recently, Begeer et al. (2014) found that ToM performances in children whose blindness involved the optic neural pathways (called ocular-plus blindness) were delayed compared to the performances of children whose blindness did not involve any neural damage (called ocular blindness). In the first case, the inability to see was due to damages of the optic trait and the brain tissue; in the second case, it was due to damages in the non-retinal parts of the eye globe. Thus, the authors argued that: "common neural mechanisms involved in visual, as well as mental, processing influence ToM development more than visual and social experience" (p. 20).

The studies presented so far essentially measured ToM competence through first-order FB tasks. Yet, ToM does not only consist in the ability to understand FB. FB might be designed to assess a specific ToM component that is more difficult to be performed in specific atypical populations, including blind children. Therefore, a more comprehensive assessment might be able to shine a light on the specific population's strengths and potentials, rather than on their weakness and delays. Following Wellman (1990), basic ToM competence is made of five components: recognizing emotions, making a distinction between physical and mental entities, appreciating that perception leads to knowledge, understanding how desires and first-order beliefs affect behavior. The individual components of ToM may display different developmental trends, and comprehensive tests are then necessary to assess ToM components both simultaneously and more precisely. Few comprehensive instruments are available: the ToM Tasks (Wellman et al. 2011), the ToM-Test (Muris et al. 1999) and ToM Storybooks (Blijd-Hoogewys et al. 2008) are among them. The ToM Tasks consists of seven components hierarchically ordered from the easier ones (diverse desire, diverse belief, knowledge access) to the more difficult ones (content FB, explicit FB, belief emotion, real-apparent emotion). The ToM-Test is designed for 5–12 year-old children and includes tasks about precursors of ToM (perception and imitation, emotion recognition, pretense and physical-reality distinction), basic ToM (belief and FB reasoning) and advanced ToM (second-order belief understanding, understanding of complex humour); children's performances improve with age. The ToM Storybooks are based on the comprehensive model proposed by Wellman and this is the tool that has been used to assess the children's ToM in the current paper. The ToM Storybooks includes 34 tasks that tap five components of the Wellman's model as part of a comprehensive assessment (emotion recognition, understanding of desire and beliefs, ability to distinguish between physical and mental entities, and awareness of the link between perception and knowledge). It has been developed in the Netherlands (Blijd-Hoogewys et al. 2008) and validated on 681 Italian children from 3 to 11 years (Molina and Bulgarelli 2012; Bulgarelli et al. 2015). The test consists of six picture books telling the story of Sam and his family; each book presents five or six tasks assessing one or more ToM components.

The ToM Storybooks were chosen for several reasons. First, it was designed according to the dynamic system approach, which aim to understand the mechanisms that shape the developmental process in a specific individual; non-linear development is expected (Thelen and Smith 1994; van Geert 2003; Blijd-Hoogewys and van Geert 2017). In this theoretical context, multiple tasks are required to measure psychological components, in order to reduce standard errors and make measurements more precise and stable (Hughes et al. 2000). Second, the test comprised both yes/no and justifications questions, that allow to assess whether children spontaneously refer to mental states when explaining people's behavior. Third, the test is suitable for children from 3 to 8 years; and finally, the Dutch version proved to be reliable and valid, especially with atypical populations (for details, see Blijd-Hoogewys et al. 2008).

The aim of the current paper was to analyze ToM development in VI children, using data collected though the ToM Storybooks, a comprehensive test that could provide more reliable and stable measurement compared to FB tasks. Moreover, to our knowledge, this was the first research run on Italian speaking blind children. It is worth noting that, although a basic universal developmental pattern for ToM was observed, ToM and language are inter-independent (Milligan et al. 2007) and FB processing also displayed some specificities as a function of cultural and linguistic features, both when comparisons were made between Western and Eastern countries, and among Western European countries belonging to two cultural Latin and Germanic clusters (Hughes et al. 2014; Molina et al. 2014; Wellman et al. 2001, 2018). Moreover, Brambring and Asbrock (2010) and Begeer et al. (2014) reported that German and Dutch blind children acquired the ability to correctly pass FB tasks at the same age. Since the research on ToM development of visually impaired children published today is mainly referred to English- or Germanic-speaking children, and it is mainly based on first-order FB tasks, the current paper contributes to filling a gap in the literature. Study 1 reports the validity of the audio version of the comprehensive test ToM Storybooks; Study 2 focuses on ToM competence in blind Italian children, assessed through the ToM Storybooks.

Study 1

Objectives

Study 1 aimed at testing the validity of the audio version of the comprehensive test ToM Storybooks, originally developed as a series of picture books. Given that standard tools usually strongly rely on visual inputs and experience, tactile and/or auditory adaptations of the tasks are necessary to allow more valid assessments of blind children.

Method

Participants

The sample encompassed 57 typically developing (TD) Italian children (28 girls and 29 boys), aged between 3 and 9 years (age in months: min = 44, max = 104; average = 74.43, SD = 18.51). The children were recruited from mainstream kindergartens or primary schools in Piedmont, Italy, in 2012.

Measures

An audio version of the ToM Storybooks has been developed by the authors of the current paper. The test is administered through a computer: the voice of a professional female speaker tells the stories, using a flat tone not to convey emotional suggestions. The experimenter stops the audio each time a question is posed to the child, and takes notes of the answers. The content of the test has been adapted in some parts: for instance, the ability to distinguish between physical and mental entities is tested through situations that do not involve sight, but only touch or sense of smell; sounds to mark the start and the end of each task and each book have also been added. One task of the test book version asks the child to choose the facial expression of emotions, picking one correct card among six (displaying happiness, sadness, anger, surprise, fear, neutral): this task is not included in the audio version of the test.

Overall, the original test encompasses 34 tasks, 75 questions that are scored dichotomously (correct or incorrect), and 18 justification questions, used to verify whether the child spontaneously refers to mental states, that are scored on a three-point scale (completely correct, partially correct, incorrect). Three scores are calculated: a "quantitative" score varying from 0 to 75, a "qualitative" score varying from 0 to 36, a "total" score obtained by summing the quantitative and qualitative scores (max. score: 111). As a whole, the audio ToM Storybooks also consists of 34 tasks, 74 yes/no questions and 18 open questions; it encompasses the quantitative, qualitative and total scores (max. score: 110).

Procedures

Both parents gave their written informed consent for the administration of the test. The children were individually assessed at their kindergarten or primary school. The audio and the book versions of the ToM Storybooks have been administered in a time span varying from 2 to 6 weeks. All the children passed the audio and the book version of the ToM Storybooks, in a randomized order: 37 children (64.9%, 21 boys and 16 girls) first passed the book version and 20 (35.1%, 8 boys and 12 girls) the audio version. The average age in months of the children who first passed the audio version did not significantly differ from the average age of the children who first passed the book version (*t* test on the age at the day of the audio administration: $t = .726$, $p = .471$; *t* test on the age at the day of the book administration: $t = .952$, $p = .345$).

Data Analysis

The Pearson's correlation was run to test the association between the children's age in months and their ToM scores, and between the scores of the audio and book versions of the test. The *t* test for dependent samples was used to check the differences between the ToM scores of the two versions of the test, and the *t* test for independent samples was used to check the differences between the ToM scores of the two groups: the children who first passed the audio version versus the children who first passed the book version.

Results and Discussion

No significant differences were detected between the scores of children who first passed the audio version and the scores of the children who first passed the book version (see Table 1).

The ToM Storybooks scores of the book and audio versions were highly and significantly correlated with children's age in months at the day of the assessment (Audio version:

Table 1 Comparison between the ToM Storybooks scores of the children who first passed the book version or the audio version

ToM score	Book version								Audio version							
	Total N = 57		Book first N = 37		Audio first N = 20		*t* test		Total N = 57		Book first N = 37		Audio first N = 20		*t* test	
	M	DS	M	DS	M	DS	*t*	*p*	M	DS	M	DS	M	DS	*t*	*p*
Quantitative	60.16	11.83	58.05	12.47	64.05	9.68	1.87	.067	47.21	11.55	46.62	12.76	48.30	9.11	.57	.569
Qualitative	10.72	6.69	9.49	6.63	13.00	6.35	1.94	.058	10.72	7.53	10.51	7.97	11.10	6.80	.28	.782
Total	70.88	17.78	67.54	18.29	77.05	15.38	1.98	.053	57.93	18.69	57.14	20.39	59.40	15.43	.47	.640

$r_{QUANTITATIVE}$ = .790, $p < .001$; $r_{QUALITATIVE}$ = .777, $p < .001$; r_{TOTAL} = .801, $p < .001$; Book version: $r_{QUANTITATIVE}$ = .748, $p < .001$; $r_{QUALITATIVE}$ = .784, $p < .001$; r_{TOTAL} = .793, $p < .001$). The correlation between age in months and the ToM scores also characterized the performances of the normative sample of the book version of the test (Bulgarelli et al. 2015; Molina and Bulgarelli 2012).

The book and audio ToM Storybooks' scores were highly and significantly associated to each other ($r_{QUANTITATIVE}$ = .856, $p < .001$; $r_{QUALITATIVE}$ = .844, $p < .001$; r_{TOTAL} = .886, $p < .001$). The average scores of the audio test were significantly lower than the book test scores, except for the qualitative score, suggesting that the audio version is more difficult than the book version. (Quantitative: M_{AUDIO} = 47.21, DS = 11.55; M_{BOOK} = 60.16, DS = 11.83; t = − 15.54, $p < .001$; Qualitative: M_{AUDIO} = 10.72, DS = 7.53; M_{BOOK} = 10.72, DS = 6.69; t = .00, $p = 1.00$; Total: M_{AUDIO} = 57.93, DS = 18.69; M_{BOOK} = 70.88, DS = 17.78; t = − 11.15, $p < .001$).

The study showed the validity of the audio version of the ToM Storybooks: both the original and the audio version were similarly associated with age, and their correlation with age was comparable to that of the Italian sample (Molina and Bulgarelli 2012). The audio version of the test proved to be a bit more difficult than the original one; nevertheless, the two versions of the test highly correlated to each other, indicating that the audio version was reliable and usable.

Study 2

Objective

Study 2 aimed at:

(1) cross-sectionally assessing visually impaired children's ToM competence through the comprehensive test ToM Storybooks, and not only through first order FB tasks, in order to verify whether the presence of possible delays in ToM development is maintained or not;
(2) comparing the results by the ToM Storybooks with the results obtained through a first order FB task included in the test in order to compare the results with previous studies in the field that showed delayed performances in VI children.

Thus, the main hypotheses are:

(1) According to the literature, based on the use of comprehensive tools, we do not expect a delay in blind children compared to sighted ones by using the Tom Storybooks;
(2) We expect to confirm the ToM delay when using the FB tasks, according to previous studies.[1]

Method

Participants

The sample of visually impaired children was retrieved from the Fondazione Robert Hollman; blind children were attending mainstream schools or preschools in Italy and they could benefit from an early intervention involving their parents/caregivers provided by the Fondazione Robert Hollman itself. Inclusion criteria were: (a) age ranged between 4 and 10 years; (b) diagnosis of congenital visual impairment; (c) no developmental or educational problems and no learning difficulties according to the neuropsychiatrist's evaluation; (d) visual acuity under 1/20, measured with the Teller's cards (Teller et al. 2005); (e) absence of diagnosis or symptoms of Autism Spectrum Disorder, assessed through the Childhood Autism Rating Scale–CARS (Schopler et al. 1980). Twenty-three children were initially involved in the study, and six were eventually excluded: two children's verbal IQ indicated a possible cognitive delay; one child's CARS score indicated the possible presence of mild to moderate autistic symptoms;

[1] A previous study showed that the items of the original version of ToM Storybooks loaded on four factors (Bulgarelli et al. 2015). Such factorial structure is not tested neither in atypical populations nor in the audio version of the test, thus sub-measures of the ToM factors are not reported in the current study.

Table 2 Blind children's characteristics

Child	Age in months	Gender	Visual impairment	Visual acuity in snellen*	CARS score
10	49.56	M	Leber congenital amaurosis	EE 1/300	20.0
23	50.00	F	Bilateral retinopathy of prematurity stage V	Absolute blindness	21.5
14	52.13	M	Bilateral congenital retinal dystrophy, Leber congenital amaurosis (suspected)	RE 1/25 LE 1/30	20.5
12	52.59	M	Bilateral retinopathy of prematurity stage V	Absolute blindness	24.0
6	61.10	M	Bilateral retinal dystrophy, optic nerve atrophy, microphthalmia, bilateral coloboma	Absolute blindness	21.5
11	62.75	M	Norrie's disease (suspected)	Absolute blindness	21.5
20	65.79	M	Chiasmatic-hypothalamic glioma	EE 1/30	19.5
17	76.92	F	Bilateral retinopathy of prematurity stage IV	EE light perception	16.5
21	81.10	F	Leber congenital amaurosis	EE 1/20	19.0
13	93.89	F	Visual impairment	RE light perception LE 1/30	16.5
3	101.23	F	Leber congenital amaurosis	Absolute blindness	18.5
19	105.26	M	Bilateral retinopathy of prematurity stage V	Absolute blindness	20.5
5	105.56	F	Bilateral microphthalmia, bilateral coloboma, persistent hyperplastic primary vitreous	Absolute blindness	15.0
7	107.59	M	Leber congenital amaurosis	Absolute blindness	19.0
22	111.16	M	Multiple congenital ocular malformations, bilateral blindness	Absolute blindness	17.0
18	122.33	F	Bilateral congenital glaucoma, malformation of the cornea	RE 1/300 LE 1/200	18.0
9	127.79	F	Bilateral congenital glaucoma	RE 1/100 LE off	16.0

**RE* right eye, *LE* left eye, *EE* both eyes

one child showed low IQ and high CARS score; two children did not complete the whole set of tests.

The final sample comprised 17 children with visual impairment or blindness (see Table 2) and 17 typically-developing children, recruited from mainstream kindergartens and schools. The two groups were matched by gender and chronological age, tolerating an age difference of maximum 6 months; no significant difference was observed between the two groups age in months ($M_{VISUALLY\_IMPAIRED}=84.22$, $DS=26.99$; $M_{CONTROL}=85.57$, $DS=25.91$; $t=-1.73$, $p=.103$). The children's verbal intelligence was measured through the verbal subtests of the Italian version of the Wechsler Intelligence Scale for Children Revised Edition (1967). The two groups showed a difference of maximum 28 scores of verbal IQ, half in negative (control group scores higher than VI children scores) and half in positive direction (VI children scores higher than control group scores); no significant difference was observed between the two groups' verbal IQ ($M_{VISUALLY\_IMPAIRED}=106.24$, $DS=11.99$; $M_{CONTROL}=106.71$, $DS=11.28$; $t=-.12$, $p=.907$).

Measures

The audio version of the ToM Storybooks was used to assess ToM competence in both the children with visual impairment and in the control group (see Study 1 for a description of the test). The ToM quantitative, qualitative and total scores were computed. The test also encompasses a first-order false belief task, namely a version of the unexpected change task, shaped on the Sally and Ann story: *Grandpa and grandma are paying Sam a visit. Sam gets rollerblades from grandpa and grandma. He's very happy with the present. 'Thank you, grandpa and grandma.' Sam puts the rollerblades in the toy trunk. Then, he goes upstairs to play with his crane. When Sam has left, his sister goes to the toy trunk. She likes to tease her brother. Lotjie hides the rollerblades in the box! And then, she goes quickly outside. Then, Sam is coming back. He wants to roller skate. Question 1: Where will Sam look for his rollerblades? Question 2: Why is Sam looking … [there]? Question 3: Where does Sam thinks his rollerblades are? Question 4: Where are they really?* Two dichotomous scores (correct or incorrect) were calculated: to "pass" the task, the child had to correctly answer to questions 1 and 4; to correctly "explain" the task, in the answers to Question 2, the child had to refer to mental states or to the fact that Laura changed the location of the roller-blades. Question 3 was a control of Question 2.

Table 3 Avarage ToM scores and standard deviations of visually impaired children and control group

ToM score	Visually impaired children (N = 17)		Control group (N = 17)		*t* test	
	M	DS	M	DS	t	*P*
Quantitative	43.65	12.67	44.65	8.76	− .42	.680
Qualitative	9.47	8.31	11.94	6.85	− 1.40	.182
Total	53.12	20.41	56.59	15.24	− .90	.380

Procedure

Both parents gave their written informed consent for the administration of the tests. The ToM Storybooks and the WISC-R/WPPSI were individually administered by the first author, in a quiet room. The CARS was filled by the first author, at the end of the assessment. Data were collected in 2010.

Data Analysis

The Pearson's correlation was run to test the association between the children's age in months and their ToM Storybooks scores, and between the visually impaired group's and control group's ToM scores. The *t* test for dependent samples was used to check the differences between the ToM scores and the False Belief Task scores of the blind children and the control group.

Results and Discussion

ToM Storybooks

The ToM Storybooks scores of the children with visual impairment highly and significantly correlated with their age in months ($r_{QUANTITATIVE} = .856, p < .001$; $r_{QUALITATIVE} = .813, p < .001$; $r_{TOTAL} = .862, p < .001$); the association of the ToM scores with the verbal IQ is also strong, ($r_{QUANTITATIVE} = .501, p = .040$; $r_{QUALITATIVE} = .415, p = .098$; $r_{TOTAL} = .480, p = .051$). The same result was obtained with the control group's scores and their age in months ($r_{QUANTITATIVE} = .697, p = .002$; $r_{QUALITATIVE} = .704, p = .002$; $r_{TOTAL} = .717, p = .001$), whereas their ToM scores were lowly associated to the verbal IQ ($r_{QUANTITATIVE} = .180, p = .490$; $r_{QUALITATIVE} = .328, p = .199$; $r_{TOTAL} = .251, p = .332$). The correlation between age and the ToM scores also characterized the performances of the normative sample of the test book version (Bulgarelli et al. 2015; Molina and Bulgarelli 2012).

The test scores of the visually impaired children and the control group did not significantly differ (see Table 3) and they strongly and significantly correlated with each other ($r_{QUANTITATIVE} = .635, p = .006$; $r_{QUALITATIVE} = .551, p = .022$; $r_{TOTAL} = .639, p = .006$).

The visually impaired children's ToM scores negatively correlated with the CARS scores ($r_{QUANTITATIVE} = -.796, p < .001$; $r_{QUALITATIVE} = -.687, p = .002$; $r_{TOTAL} = -.773, p < .001$). This result showed that the lower presence of autistic traits was correlated with higher ToM performances. In fact, the convergent validity of the Dutch test was tested in samples of children with Pervasive Developmental Disorders and typically developing children (Blijd-Hoogewys et al. 2008): their scores negatively correlated with the scores of the Vineland Adaptive Behavior Scales scores (Sparrow et al. 1984) and the Children's Social Behaviour Questionnaire (Luteijn et al. 2000).

First-Order False Belief Task

Differences between the visually impaired children's and the control group children's scores in the unexpected change task were analysed with the t-test. The two groups performed at the same level (Pass: $M_{VISUALLYIMPAIRED} = .35$, DS = .49, $M_{CONTROL} = .35$, DS = .49, t = .00, $p = 1.00$; Explain: $M_{VISUALLYIMPAIRED} = .47$, DS = .72, $M_{CONTROL} = .29$, DS = .47, t = .90, $p = .382$). It is worth noting that the correct response rate was low for both samples.

General Discussion

The current paper had the major objective to contribute to the knowledge about VI children's ToM competence and development, thanks to the use of a comprehensive test, the ToM Storybooks, that could allow more precise, stable and reliable measurement. The original book version of the test had to be adapted for blind subjects; thus, Study 1 demonstrated the validity of the test audio version. Moreover, in Study 2, the ToM Storybooks scores were negatively correlated with the CARS scores, further supporting the construct validity of the ToM Storybooks. Study 1 also showed that, for typically developing children, the audio test was more difficult than the book version.

Study 2 focused on VI children's ToM competence and development and showed that they performed very similarly to TD children matched on chronological age and gender. In fact, accordingly to hypothesis 1, the mean scores of the ToM Storybooks did not differ between the two groups. This result was in line with the recent work by Pijnacker et al. (2012): in both studies, narrative tasks were used to assess ToM and one possible explanation of this result is that, compared to tactile type of tasks, auditory tasks were more suitable for blind children. Another indicator of typical development of ToM competence lies in the fact that the blind children's scores highly and significantly correlated with their chronological age in months and verbal IQ: this result was usually observed in typical development (Blijd-Hoogewys et al. 2008; Bulgarelli et al. 2015; Molina and Bulgarelli 2012) and not observed in children with autistic spectrum disorder (Molina et al. 2010). The correlation of ToM scores with chronological age was also reported in Pijnacker et al. (2012) and Roch-Levecq (2006). Study 2 showed no differences between the VI and TD children's ToM scores and verbal IQ, whilst a strong association between ToM scores and verbal IQ was observed within the VI children, and a weaker association was observed within the TD children. This result suggests that language could support an adaptive compensation of the lacking visual experience in building VI children's ToM competence. Also, it is possible that VI children mainly rely on language understand everyday experience, more than TD children do, and consequently, language had a stronger effect on their ToM performance. This pattern of results is in line with Fraiberg (1977) and Pérez-Pereira and Conti-Ramsden (1999), who argued that blind children's different experience of the world ended up in developing competences similarly to typical peers, once supported in developing their strengths based on other senses. Usually, blind children develop language quite easily and irregularities and delays shown in the early years are usually overcome by school age (Tadić et al. 2010).

Differently from what expected by hypothesis 2 and the previous studies in field (Begeer et al. 2014; Brambring and Asbrock 2010; Green 2004; McAlpine and Moore 1995; Minter et al. 1998; Peterson et al. 2000; Roch-Levecq 2006), the VI and TD children showed also similar performances on the first order FB task included in the ToM Storybooks. Thus, the results reported in the current paper differ from those that showed delayed ToM development when first order FB tasks are used to assess blind children's competence. Nevertheless, it should be taken into account that the audio version of the ToM Storybooks showed to be more difficult than the book version for the typically developing children involved in Study 1: this could have disadvantaged the control group of Study 2 and reduced the gap between the blind children and their matched peers. In fact, as discussed above, the ToM scores of the control group of Study 2 were not correlated with their verbal IQ, and this was an unexpected result. As Pijnacker et al. (2012) stated, the processing of auditory tasks might place a high demand on working memory capacity, to keep in mind the story information for sufficient time to answer, and blind children seemed to have a better short-term memory for auditory verbal information (Swanson and Luxenberg 2009).

This picture brings along two considerations. On one side, more research is needed to understand developmental ToM pathways in typical and atypical development. A possible future area of research could compare VI children and children with autism matched on verbal IQ in order to further understand the role of language on ToM development and on children's development. Also, measures related to children's conversational background and parents' ToM abilities should be collected as well, to better test the hypothesis of the role of conversation on mind comprehension. On the other side, the use of standardized comprehensive tests is crucial to assess children's competences and development: the possibility to evaluate ToM components through a set of tasks enables more precise and stable measurement, especially with atypical populations (Hughes et al. 2000). It is also worth noting that the current paper presents the first results collected through a comprehensive test with Italian blind children. Cultural influence on ToM were retrieved and, so far, typically developing British children were found to out-perform Italian children ToM tasks (Hughes et al. 2014), whereas Italian children obtained higher scores on the ability to hide emotions compared to German children (Molina et al. 2014). Cross-cultural comparisons among atypically developing children are needed to confirm these first results. With respect to ToM development in blind children, crucial data could also be detected by comparing subjects with different types of visual impairment: those that involve neural systems and those that are solely located in the eye globe, as done in Begeer et al. (2014). In fact, potentially different developmental trends could characterize children whose cause of blindness is located in brain areas and children whose cause of impairment is not. This could be an area of further development: this analysis was not possible in our study given the limited number of subjects in the two subgroups considered (ocular versus ocular plus).

Acknowledgments Study 1 was run by the second and third authors, who would like to thank Maria Nives Sala and Nicoletta Pinto for their support in data collection. Study 2 reported part of the results of the first author's Ph.D. Dissertation Thesis; she would like to thank Prof. Adriano Pagnin, PhD supervisor from the University of Pavia; Roberta Ceroni for the contribution in data collection; Dr.ssa Paola Caldironi, Executive Director of the Fondazione Robert Hollman at the time of the research, and Josee Lanners, Deputy Director of the Cannero Riviera Centre. We thank the children and the families who participated in the studies, the Fondazione Robert Hollman of Cannero Riviera (Italy) and Padova (Italy), and the schools that were involved in the research project.

Author Contributions GB conceived of the Study 2, participated in its design and coordination, performed its measurement and helped to draft the manuscript; DB conceived of the Study 1 and participated in its design and coordination, performed its measurement and the statistical analysis of both studies, and drafted the manuscript; PM conceived of the Study 1, participated in its design and coordination, helped to perform the analyses and to draft the manuscript. All authors read and approved the final manuscript.

Compliance with Ethical Standards

Conflict of interest The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Ethical Approval This study was carried out in accordance with the ethical recommendations of the Ethical Code of the Italian Psychologists Association. All subjects gave their written informed consent in accordance to the Declaration of Helsinki.

References

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

Bartoli, G., Soni, M., Montefusco, L., Lanners, J., Di Nunzio, S., & Cegalin, M. (2010). *"Autistic like features" in blind children. The challenge of understanding their meaning in the development process*. Poster presented at RNIB Conference 'Children who have visual impairment and autism', London (UK).

Bedny, M., Pascual-Leone, A., & Saxe, R. R. (2009). Growing up blind does not change the neural bases of Theory of Mind. *Proceedings of the National Academy of Sciences, 106*(27), 11312–11317. https://doi.org/10.1073/pnas.0900010106.

Begeer, S., Dik, M., Asbrock, D., Brambring, M., & Kef, S. (2014). A new look at theory of mind in children with ocular and ocular-plus congenital blindness. *Journal of Visual Impairment and Blindness, 108*(1), 17–27. https://doi.org/10.1177/0145482X1410800103.

Blijd-Hoogewys, E., & van Geert, P. L. (2017). Non-linearities in theory-of-mind development. *Frontiers in Psychology, 7,* 1970. https://doi.org/10.3389/fpsyg.2016.01970.

Blijd-Hoogewys, E. M. A., van Geert, P. L. C., Serra, M., & Minderaa, R. B. (2008). Measuring theory of mind in children: Psychometric properties of the tom storybooks. *Journal of Autism and Developmental Disorders, 38*(10), 1907–1930. https://doi.org/10.1007/s10803-008-0585-3.

Brambring, M., & Asbrock, D. (2010). Validity of false belief tasks in blind children. *Journal of Autism and Developmental Disorders, 40*(12), 1471–1484. https://doi.org/10.1007/s10803-010-1002-2.

Bulgarelli, D., & Molina, P. (2013). Teoria della Mente nei bambini con Disturbo Specifico di linguaggio: Una questione di competenza o di performance? [Theory of Mind in children with Specific Language Impairment: A matter of competence or performance?]. *Giornale Italiano di Psicologia, 60*(4), 761–784. https://doi.org/10.1421/76946.

Bulgarelli, D., Testa, S., & Molina, P. (2015). Factorial structure of the 'ToM Storybooks': A test evaluating multiple components of theory of mind. *British Journal of Developmental Psychology, 33*(2), 187–202. https://doi.org/10.1111/bjdp.12062.

Farrenkopf, C., & Davidson, I. F. W. K. (1992). The development of perspective-taking abilities in young blind children. *RE: View, 24*(1), 7–23.

Flavell, J. H. (2004). Theory-of-mind development: Retrospect and prospect. *Merrill-Palmer Quarterly, 50,* 274–290.

Fraiberg, S. (1977). *Insights from the blind. Comparative studies of blind and sighted infants*. New York: Basic Books.

Gergely, G., & Watson, J. S. (1996). The social biofeedback model of parental affect-mirroring. *The International journal of psychoanalysis, 77*(6), 1181.

Green, S., Pring, L., & Swettenham, J. (2004). An investigation of first-order false belief understanding of children with congenital profound visual impairment. *British Journal of Developmental Psychology, 22*(1), 1–17. https://doi.org/10.1348/026151004772901087.

Hobson, R. P. (1990). On acquiring knowledge about people and the capacity to pretend: Response to Leslie (1987). *Psychological Review, 97*(1), 114–121. https://doi.org/10.1037/0033-295X.97.1.114.

Hobson, P. (2004). *The cradle of thought: Exploring the origins of thinking*. Oxford: Oxford University Press.

Hughes, C., Adlam, A., Happé, F., Jackson, J., Taylor, A., & Caspi, A. (2000). Good test-retest reliability or standard and advanced false belief tasks across a wide range of abilities. *Journal of Child Psychology and Psychiatry, 41*(4), 483–490. https://doi.org/10.1111/1469-7610.00633.

Hughes, C., Devine, R. T., Ensor, R., Koyasu, M., Mizokawa, A., & Lecce, S. (2014). Lost in translation? Comparing British, Japanese, and Italian children's theory-of-mind performance. *Child Development Research, 2014,* 1–10. https://doi.org/10.1155/2014/893492.

Luteijn, E., Luteijn, F., Jackson, S., Volkmar, F., & Minderaa, R. B. (2000). The children's social understanding questionnaire for milder variants of PDD problems: Evaluation of the psychometric characteristics. *Journal of Autism and Developmental Disorders, 30*(4), 317–330. https://doi.org/10.1023/A:1005527300247.

McAlpine, L. M., & Moore, C. L. (1995). The development of social understanding in children with visual impairments. *Journal of Visual Impairment and Blindness, 89*(4), 349–358.

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

Minter, M., Hobson, R. P., & Bishop, M. (1998). Congenital visual impairment and 'theory of mind'. *British Journal of Developmental Psychology, 16*(2), 183–196. https://doi.org/10.1111/j.2044-835X.1998.tb00918.x.

Mitchell, P. (1997). *Introduction to theory of mind. children, autism and apes*. London: Arnold.

Molina, P., & Bulgarelli, D. (2012). La standardizzazione italiana del test ToM Storybooks: Dati preliminari [The Italian standardization of the test ToM Storybooks: Prelimanary data]. *Giornale Italiano di Psicologia, 39*(4), 863–880. https://doi.org/10.1421/73146.

Molina, P., Bulgarelli, D., Henning, A., & Aschersleben, G. (2014). Emotion nderstanding: A cross-cultural comparison between Italian and German preschoolers. *European Journal of Developmental Psychology, 11*(5), 592–607. https://doi.org/10.1080/17405629.2014.890585.

Molina, P., Bulgarelli, D., & Salomone, E. (2010). *Discriminant validity of the ToM Storybooks and the TEC in an Italian sample of PDD children*. Poster presented at the IX International Congress Autism Europe, Catania (I).

Muris, P., Steerneman, P., Meesters, C., Merckelbach, H., Horselenberg, R., van den Hogen, T., et al. (1999). The TOM test: A new instrument for assessing theory of mind in normal children and children with pervasive developmental disorders. *Journal of Autism and Developmental Disorders, 29*(1), 67–80. https://doi.org/10.1023/A:1025922717020.

Perez-Pereira, M., & Conti-Ramsden, G. (1999). *Language development and social interaction in blind children*. Hove: Psychology Press.

Pérez-Pereira, M., & Conti-Ramsden, G. (2005). Do blind children show autistic features. In L. Pring (Ed.), *Autism and blindness: Research and reflections* (pp. 99–127). London: Whurr.

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

Perner, J., Leekam, S. R., & Wimmer, H. (1987). Three-year-olds' difficulty with false belief: The case for a conceptual deficit. *British Journal of Developmental Psychology, 5,* 125–137. https://doi.org/10.1111/j.2044-835X.1987.tb01048.x.

Peterson, C. C. (2009). Development of social-cognitive and communication skills in children born deaf. *Scandinavian Journal of Psychology, 50*(5), 475–483. https://doi.org/10.1111/j.1467-9450.2009.00750.x.

Peterson, C. C., & Wellman, H. M. (2018). Longitudinal theory of mind (ToM) development from preschool to adolescence with and without ToM delay. *Child Development*. https://doi.org/10.1111/cdev.13064.

Peterson, C. C., Peterson, J. L., & Webb, J. (2000). Factors influencing the development of a theory of mind in blind children. *British Journal of Developmental Psychology, 18*(3), 431–447. https://doi.org/10.1348/026151000165788.

Pijnacker, J., Vervloed, M. P., & Steenbergen, B. (2012). Pragmatic abilities in children with congenital visual impairment: An exploration of non-literal language and advanced theory of mind understanding. *Journal of Autism and Developmental Disorders, 42*(11), 2440–2449. https://doi.org/10.1007/s10803-012-1500-5.

Roch-Levecq, A. C. (2006). Production of basic emotions by children with congenital blindness: Evidence for the embodiment of theory of mind. *British Journal of Developmental Psychology, 24*(3), 507–528. https://doi.org/10.1348/026151005X50663.

Scheeren, A. M., de Rosnay, M., Koot, H. M., & Begeer, S. (2013). Rethinking theory of mind in high-functioning autism spectrum disorder. *Journal of Child Psychology and Psychiatry, 54*(6), 628–635. https://doi.org/10.1111/jcpp.12007.

Schopler, E., Reichler, R. J., DeVellis, R. F., & Daly, K. (1980). Toward objective classification of childhood autism: Childhood autism rating scale (CARS). *Journal of Autism and Developmental Disorders, 10*(1), 91–103. https://doi.org/10.1007/BF02408436.

Shields, J., Varley, R., Broks, P., & Simpson, A. (1996). Social cognition in developmental language disorders and high-level autism. *Developmental Medicine and Child Neurology, 38*(6), 487–495. https://doi.org/10.1111/j.1469-8749.1996.tb12109.x.

Sparrow, S. S., Balla, D. A., & Cicchetti, D. V. (1984). *Vineland Adaptive Behaviour Scales*. Circle Pines: American Guidance Services.

Sullivan, K., Zaitchik, D., & Tager-Flusberg, H. (1994). Preschoolers can attribute second-order beliefs. *Developmental Psychology, 30*(3), 395–402. https://doi.org/10.1037/0012-1649.30.3.395.

Swanson, H. L., & Luxenberg, D. (2009). Short-term memory and working memory in children with blindness: Support for a domain general or domain specific system? *Child Neuropsychology, 15*(3), 280–294. https://doi.org/10.1080/09297040802524206.

Tadić, V., Pring, L., & Dale, N. (2010). Are language and social communication intact in children with congenital visual impairment at school age? *Journal of Child Psychology and Psychiatry, 51*(6), 696–705. https://doi.org/10.1348/026151008X310210.

Teller, D. Y., Dobson, V., & Mayer, D. L. (2005). *Teller Acuity Cards II. TAC II. Handbook*. Chicago: Stereo Optical Co., Inc.

Thelen, E., & Smith, L. B. (1994). *A dynamic systems approach to the development of cognition and action*. Cambridge: The MIT Press.

van Geert, P. L. C. (2003). Dynamic Systems Approaches and Modeling of Developmental Processes. In J. Valsiner & K. J. Conolly (Eds.), *Handbook of developmental psychology* (pp. 640–672). London: Sage.

Wechsler, D. (1967) *Wechsler intelligence scale for children revised.* New York, NY: Psychological Corporation. Trad It: Rubini, V., & Padovani, F. (1986) (Eds.) *WISC-R Scala d'Intelligenza Wechsler per Bambini Riveduta.* Firenze, I: Organizzazioni Speciali.

Wellman, H. M. (1990). *The child's theory of mind*. Cambridge: MIT Press.

Wellman, H. M. (2018). Theory of mind: The state of the art. *European Journal of Developmental Psychology, 15,* 1–28. https://doi.org/10.1080/17405629.2018.1435413.

Wellman, H. M., Cross, D., & Watson, J. (2001). Meta-analysis of theory-of-mind development: The truth about false belief. *Child Development, 72*(3), 655–684. https://doi.org/10.1111/1467-8624.00304.

Wellman, H. M., Fang, F., & Peterson, C. C. (2011). Sequential progressions in a theory-of-mind scale: Longitudinal perspectives. *Child Development, 82*(3), 780–792. https://doi.org/10.1111/j.1467-8624.2011.01583.x/.

Williams, M. E., Fink, C., Zamora, I., & Borchert, M. (2014). Autism assessment in children with optic nerve hypoplasia and other vision impairments. *Developmental Medicine and Child Neurology, 56*(1), 66–72. https://doi.org/10.1111/dmcn.12264.

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

Yirmiya, N., Erel, O., Shaked, M., & Solomonica-Levi, D. (1998). Meta-analyses comparing theory of mind abilities of individuals with autism, individuals with mental retardation, and normally developing individuals. *Psychological Bulletin, 124*(3), 283–307. https://doi.org/10.1037/0033-2909.124.3.283.