



The Psychometric Properties of the Children’s Social Understanding Scale among Atypically Developing Children and Adolescents

Burcu Besiroglu ¹ · Ayca Alayli ^{2,3} · Deniz Tahiroglu ⁴ · Bilge Selcuk ⁵

Accepted: 13 January 2022 / Published online: 26 January 2022

© The Author(s), under exclusive licence to Springer Science+Business Media, LLC, part of Springer Nature 2022

Abstract

Theory of mind (ToM) has been frequently assessed via behavioral assessments. Recently, Tahiroglu and colleagues developed the parent-report Children’s Social Understanding Scale (CSUS) to assess individual differences in ToM. They found it reliable and valid in assessing ToM of typically developing preschoolers. In the present study, we examined the psychometric properties of the CSUS in ToM assessment of children and adolescents with atypical development, who are known to have difficulties in understanding others’ mental states. In two studies, we collected data from 106 children and adolescents with autism spectrum disorder (ASD; $M_{age} = 12.06$ years, $SD = 2.91$) and 70 children and adolescents with hearing impairment (HI; $M_{age} = 7.3$ years, $SD = 2.3$). The children’s social understanding was assessed via the CSUS (short form in the ASD and full form in the HI samples) and behavioral ToM tasks (tasks with low-verbal demands and standard ToM tasks). Receptive language in both samples and nonverbal intelligence in ASD sample were also assessed. Analyses revealed high internal consistency for the CSUS in both samples. As expected, the CSUS correlated with performance on behavioral ToM tasks, receptive language, and nonverbal intelligence. The CSUS continued to be significantly associated with performance on behavioral ToM tasks after controlling for receptive language and nonverbal intelligence in ASD sample, and controlling for age and receptive language in HI sample. These results demonstrate that the CSUS is a reliable and valid tool in measuring ToM among children and adolescents with ASD and those with HI in Turkey.

Keywords Children’s Social Understanding Scale · Theory of mind · Parent report · Autism spectrum disorder · Hearing impairment

Highlights

- Research examining theory of mind has mostly focused on behavioral assessments.
- Parent-reports would be beneficial for assessing theory of mind especially in atypical samples.
- The Children’s Social Understanding Scale is used in different cultures, yet, not widely studied in atypical samples.
- The scale has good psychometric properties in samples of individuals with autism spectrum disorder and hearing impairment.
- The scale captures individual differences in theory of mind of individuals with autism spectrum disorder and hearing impairment.

Theory of mind (ToM) refers to the ability to understand the mental states of oneself and others to predict and explain

behavior (Wellman, 2014). It is considered as one of the fundamental milestones in socio-cognitive development. It

✉ Burcu Besiroglu
burcu.besiroglu@bilgi.edu.tr

¹ Department of Psychology, Istanbul Bilgi University, Istanbul, Turkey

² Department of Sociology, University of Amsterdam, Amsterdam, The Netherlands

³ Department of Child and Family Studies, Utrecht University, Utrecht, The Netherlands

⁴ Department of Psychology, Bogazici University, Istanbul, Turkey

⁵ Present address: Department of Psychology, Koc University, Istanbul, Turkey

helps children to understand the social world, coordinate their thoughts, engage in complex social behaviors (Astington, 2003, Wellman, 2014), and is linked with pro-social behaviors (Imuta et al., 2016), social competence (Astington, 2003), empathy, morality (Lonigro et al., 2013), and better academic performance (Clemmensen et al., 2018). Delay or deficit in this ability leads to difficulties in relationships with others and social adjustment (Nader-Grosbois et al., 2013). A deficit in ToM is one of the defining characteristics of individuals with autism spectrum disorder (ASD; Baron-Cohen et al., 1985) and there are other atypically developing groups (e.g., individuals with hearing impairment, HI) who are known to have delays or deficits in understanding others' mental states (Wellman & Peterson, 2013).

ToM has been predominantly measured via behavioral assessments which depend on participants' performance (e.g., Brent et al., 2004, Peterson et al., 2007). More specifically, research on ToM has mostly focused on children's false belief understanding, which refers to children's ability to understand that others can have a belief that contradicts reality (Wellman et al., 2001). In the past decade, researchers started to distinguish between explicit and implicit ToM performance. Children's explicit ToM, which is defined as the ability to provide reasons for explaining their and others' behaviors via mental states, has been assessed via standard ToM tasks, in which children are given a story with multiple characters and are directly asked about what they expect the story characters believe and/or how they will act (Wellman et al., 2001). These studies have revealed that typically developing children start to pass ToM tasks such as false belief understanding around the age of 4 years (Wellman, 2014, Yirmiya et al., 1998). Yet, some researchers have criticized explicit ToM tasks on grounds that they conceal younger children's ToM abilities due to high executive function and linguistic demands (Baillargeon et al., 2010). These critiques led to the emergence of implicit ToM tasks, through which children's ToM abilities are inferred via violation-of-expectation and eye-tracking tasks. These implicit ToM tasks have indeed demonstrated that ToM develops earlier than was detected by the explicit ToM tasks (Baillargeon et al., 2010). For instance, even as young as 7-month-old infants act as if they have ToM when assessed with implicit ToM tasks (Kovács et al., 2010).

Yet, children with ASD and children with HI who have hearing parents are known to pass both explicit (Happé & Frith, 1996, Hoogenhout & Malcolm-Smith, 2014 for ASD; Peterson & Wellman, 2009, Peterson et al., 2012 for HI) and implicit ToM tasks (Senju et al., 2010 for ASD; de Villiers & de Villiers, 2012 for HI) at a later chronological and (verbal) mental age. According to Happé (1995), children with ASD, on average, have a verbal mental age of a 9-year-old to pass ToM tasks. The youngest mental age to

pass such tests is found to be 5.6 years, while the youngest chronological age is 14.10 years. These findings demonstrate a significant delay in ToM development of children with ASD in comparison to typical development. For individuals with HI, it is not yet clear how much of a delay in ToM development to expect due to the variety of sample characteristics (e.g., additional sign language input, mode of schooling, and communication); however, there is a growing consensus that typical development can be reached with early device use (e.g., starting earlier than 12 months) (Leigh et al., 2013). Literature on ASD and HI has shown that deficits in ToM are related to key impairments in the social domain (Peterson, 2003, Tager-Flusberg, 2007). Thus, it is important to accurately assess this ability to better understand children's socio-cognitive development, both in typically and atypically developing groups.

In addition to behavioral assessments, there are several other-informant assessment tools available, which are recently added to the repertoire of measures to be used with typically developing children as well as atypically developing groups. The main goal of this current research is to test the psychometric properties of one of these new measures, the parent-report Children's Social Understanding Scale (CSUS, Tahiroglu et al., 2014) in two atypically developing groups: children and adolescents with ASD and HI.

Limitations of Current Behavioral Measures of Theory of Mind

ToM is a complex and multifaceted construct, including abilities such as distinguishing appearance from reality, perspective-taking, and understanding differences between one's own and others' beliefs, desires, and emotions. For decades, research on ToM of both typically developing children and children with ASD or HI has been dominated by behavioral assessments (Wellman, 2014). Researchers have preferred behavioral assessments over self or other-informant reports since these tasks enable direct observation of children's ToM performance without the potential drawbacks of informant reports, such as parents not being able to report accurately and overestimating or underestimating their children's capacity.

However, researchers have been raising concerns over the use of behavioral measures of ToM as the main assessment tool. One problem is the focus of research on tasks such as false belief understanding, which limit the assessment of individual differences since they categorize children as either having false belief understanding capacities or not (Roth & Leslie, 1998). In addition, due to focusing only on false belief, these tasks fall short of assessing the highly complex nature of ToM (Astington,

2001, Wellman, 2012). To compensate for the limitations of using a single task or measuring only one dimension of ToM, researchers sometimes prefer assessing and aggregating scores of tasks tapping into different mental states to address the multifaceted nature of ToM (e.g., Baron-Cohen et al., 1997, Muris et al., 1999, Peterson et al., 2005). Yet, the use of a battery of ToM tasks would require more time, effort, and attention than the use of a single task. Thus, it is not always practical, especially when assessing other developmental capacities in addition to ToM.

Relatedly, many of the behavioral ToM tasks, and false belief understanding tasks in particular, rely heavily on language and cognitive abilities such as remembering the story events and inhibiting spontaneous responses (e.g., inhibiting the predominant/salient response to give an accurate answer about the location of an object; Astington & Jenkins, 1999, San José Cáceres et al., 2014, Tager-Flusberg, 2007). It is argued that even when the tasks do not require a verbal response, children have to handle the linguistic and cognitive complexity of ToM tasks; they have to remember the instructions, understand the content of the stories, and process the questions to provide appropriate answers (Astington & Jenkins, 1999). Thus, the tasks are criticized for not being ‘pure’ in the assessment of ToM but rather the performance is very much influenced by the children’s linguistic and cognitive abilities.

Another problem raised is that children’s performance in ToM tasks generally do not show high correspondence with their performance in real-life situations that require social understanding, such as differentiating own and others’ perspectives, recognizing and understanding emotions or having an empathic stance and morality (Begeer et al., 2010, de Villiers & de Villiers 2000, Raver & Leadbeater, 1993). One way to overcome this limitation has been to use naturalistic techniques (e.g., observing participants’ everyday behaviors). These techniques have advantages over experimental measures such that the reports and observations reflect participants’ behavior in natural contexts without dependence on specific tasks demands such as being in an unfamiliar setting with a researcher (Peskin & Ardino, 2003), and thus have higher ecological validity (Begeer et al., 2010, Kleinman et al., 2001, Spek et al., 2010). However, despite being highly informative, unlike standardized measures, these naturalistic ToM measures rely on spontaneous behaviors, which might be infrequent, anecdotal, and free of control, rendering conclusions difficult.

These criticisms made against behavioral assessments of ToM are even more pronounced for children with atypical development (Tager-Flusberg, 2007). It is argued, for instance, that typically developing children use their social insight and general cognitive skills to answer the questions in ToM tasks. However, children with ASD or HI may rely

on language and other nonsocial cognitive processes more than their social insight when assessed on behavioral ToM tasks. For example, instead of intuitively answering questions like children with typical development, these children break down ToM questions into their linguistic and general logical elements. For a meaningful response, they systematically try to find clues from statements instead of giving an automatic response (Kaland et al., 2008). Also, these children have broader difficulties in understanding social interactions. When an unfamiliar person (e.g., the researcher) asks questions to them, they might get frustrated and miss some aspects of the questions which could lead to slower response time. These factors can also influence the children’s motivation, for instance by decreasing their interest level and attention and increasing fatigue (Begeer et al., 2003). Thus, it is argued that the ToM performance of children with ASD and HI is more sensitive to the problems listed above (de Villiers & de Villiers, 2000, Tager-Flusberg, 2007, Van Herwegen et al., 2013).

Parent-Report Measurements of Theory of Mind

Every measurement tool has its drawbacks and a certain degree of error. To minimize source and setting errors, it is desirable to have complementary tools (e.g., self-/other-reports in addition to behavioral and/or observational methods) to assess psychological constructs (Frick et al., 2010). Moreover, complementary measures help us better understand the phenomenon via providing information about different situations and diverse perspectives. Parents frequently serve as informants in the assessment of children’s social and cognitive abilities such as social competence (e.g., Social Competence and Behavior Evaluation-30; LaFreniere & Dumas, 1996) and executive functions (e.g., The Behavior Rating Inventory of Executive Function; Gioia et al., 2000). Parents play a critical role in children’s ToM development (Pavarini et al., 2012) and can also be a good source of information about their children’s mental states—both explicit (e.g., whether children use certain kinds of words/phrases) and implicit (e.g., whether they engage in certain kinds of behaviors that would reflect ToM), as they have a long-time opportunity to observe their children over the years and different situations with different correspondents.

Currently, there are only a few measures based on caregiver reports (e.g., ToM Behavioral Checklist [Begeer et al., 2015], Everyday Mindreading Skills and Difficulties Scale [Peterson et al., 2009], and ToM Inventory [Hutchins et al., 2012]) to assess children’s ToM-related behaviors. These measures have been designed based on interviews with parents of individuals with ASD or the literature

outlining the development of ToM in children with typical development and ASD. The reliability and validity of these measures have been tested with typically and atypically developing children and adolescents, mostly in ASD samples, between the ages of 2 and 18 years. In addition to ASD samples, the ToM Inventory was utilized in a sample with HI in a pilot study and showed good psychometric properties (Hutchins et al., 2017). In the current study, we aimed to investigate the psychometric properties of yet another parent-report ToM scale, the CSUS (Tahiroglu et al., 2014). Contrary to aforementioned caregiver reports in which questions were mostly based on atypical development, the CSUS was originally developed based on the typical development of preschool-aged children. The CSUS has been administered in different cultures and languages such as French (Brosseau-Liard & Poulin-Dubois, 2018), Polish (Białecka-Pikul & Stępień-Nycz, 2019, Smogorzewska et al., 2019), Korean (Jang & Shin, 2018), and Turkish (Ekerim-Akbulut et al., 2021) and found to have good psychometric properties. In our study, we aimed to examine how the scale's psychometric properties would fare when used in the assessment of ToM in Turkish children and adolescents with atypical development, specifically ASD and HI.

Children's Social Understanding Scale (CSUS)

Tahiroglu and colleagues (2014) developed the CSUS to extend the available ToM assessment tools to capture individual differences in typically developing children's ToM. The scale aims to assess the frequency with which the child uses mental-state terms and behaves in ways that reflect explicit and implicit ToM (see Tahiroglu et al., 2014, for detailed information). In other words, it comprises both verbalization and behavioral aspects that reflect ToM ability, unlike the other ToM scales which generally assess only one of these aspects. The CSUS also taps into the multi-faceted nature of ToM and assesses multiple different aspects of ToM, namely understanding of belief (e.g., understanding false beliefs), knowledge (e.g., understanding different levels of knowledge), perception (e.g., understanding the difference between perceptual appearances and reality), desire (e.g., understanding different desires), intention (e.g., understanding the difference between intentions and outcomes), and emotion (e.g., understanding multiple emotions about the same situation). Although the other ToM scales also investigate the complex structure of ToM, only the CSUS assesses ToM from a sequential perspective demonstrated by Wellman and Liu (2004). It consists of the multi-faceted construct of ToM, capturing incrementally complex mental states such as desires,

beliefs, intentions, and emotions. The full form has 42 items in which each aspect of ToM is assessed with 7 items. The items were selected from a large pool of items derived from ToM tasks commonly administered to children, key components of ToM as reported in the literature, children's everyday behaviors reflecting ToM understanding and interviews with parents and evaluation of items by international ToM experts. The short form of the CSUS includes 18 items, with each aspect of ToM being assessed with 3 items from the subset of 42 items of the full form based on content coverage (i.e., measuring the six aspects of ToM assessed in full form) and correlation of items with behavioral ToM tasks.

Tahiroglu et al. (2014) conducted three studies to assess the psychometric properties of the CSUS. These studies indicated good internal consistencies (α s ranging from 0.81 to 0.94) and a strong short-term test-retest reliability (r s = 0.88, p s < 0.01) for both the full and short forms of the scale, among typically developing North American preschoolers. The scale scores were validated and found to be closely related to behavioral assessments of ToM (i.e., knowledge access, contents and explicit false belief, Level 2 perspective-taking, appearance-reality, and restricted view tasks) and cognitive skills such as prospective memory, working memory, and planning.

The psychometric properties of the Turkish version of the CSUS-Short form has been recently investigated among typically developing preschoolers (Ekerim-Akbulut et al., 2021). The study confirmed one-factor model for the Turkish version of the CSUS-Short form, consistent with the original scale. It revealed that the CSUS-Short form has a good internal consistency (α s = 0.84 at both assessment points, one-year-apart) and a good test-retest reliability (r = 0.52, p < 0.001). The CSUS-Short form score was significantly correlated with Turkish children's ToM performance (i.e., knowledge access, contents false belief, and unexpected change tasks, r s ranging between 0.16 and 0.28). It was also significantly associated with their social and cognitive skills (i.e., social competency, executive functions, and receptive language), sex, and age, providing evidence for the validity of the scale in use with Turkish preschool children. The psychometric properties of the CSUS have been also studied across different cultures in school-age children with typical development as well as in populations with atypical development. The scale has been found to have good psychometric properties among Polish elementary schoolers with typical development whose ages ranged between 6 and 9 years (Smogorzewska et al., 2019). These studies demonstrate that, in typically developing populations, the scale can be used to measure ToM of preschool- and elementary school-age children. In the same study, Smogorzewska and colleagues (2019) also investigated the psychometric properties of the CSUS in

elementary schoolers with atypical development, specifically those with mild intellectual disability and HI. The study was conducted with 726 Polish parents from diverse educational backgrounds. The study had a comparatively wider age range for children and adolescents with atypical development (6–9 years for typical development, 6–12 years for mild intellectual disability, and 5.11–10.11 for HI). Analyses revealed a very high reliability (α s ranged from 0.93 to 0.96) in all samples. The CSUS had strong associations with behaviorally measured ToM (i.e., ToM scale and the Faux Pas Recognition Test) and social skills when controlling for language. The results also demonstrated that children and adolescents with atypical development (both mild intellectual disability and HI) had lower CSUS scores than typically developing children, with children with mild intellectual disability having the lowest scores. This study also showed the usefulness of the CSUS in measuring ToM of children and adolescents with atypical development whose ages ranged between 5.11 and 12 years.

There is also some preliminary evidence that the CSUS is reliable and valid when used to assess ToM in children and adolescents with ASD in North America (Tahiroglu et al., 2014). In a pilot study, Tahiroglu et al. (2014) examined the discriminant validity of the CSUS in assessing ToM of children and adolescents with ASD ($n = 15$, age range of 10–16 years) and typical development ($n = 18$). When children's and adolescents' age and intelligence were controlled for in analyses, typically developing children and adolescents were found to have higher CSUS scores than children and adolescents with ASD (matched on chronological age). Children's and adolescents' CSUS scores also significantly predicted the severity of autistic traits in both typically developing and ASD samples. Thus, the results of the pilot study provided preliminary evidence for the discriminant validity and the usefulness of the scale in children and adolescents with ASD. However, the study was conducted with a very small sample size and focused on a narrow age range. Thus, to test the validity of the CSUS in ASD, it would be beneficial to assess the psychometric properties in a larger sample.

Goals and Importance of the Study

ToM assessment instruments are important for identifying ToM difficulties and evaluating treatment progress in atypically developing children, such as children with ASD, HI, intellectual disabilities, or specific language impairment (Ahmadi et al., 2015, Blijd-Hoogewys et al., 2008). Thus, having an additional measure of ToM such as a parent-report to be used in both typically and atypically developing samples would be highly useful in extending our understanding of ToM deficits of children and adolescents with

typical and atypical development. This would not only provide us with the practical advantage of comparing different groups but also deepen our understanding of ToM development in both typically and atypically developing samples.

The use of parent reports in ToM assessment is relatively recent, and, being widely adapted and validated in different cultures, the CSUS has the potential to be a tool for use in cross-cultural comparisons as well as comparisons of typically and atypically developing groups. Although many studies on the validation of the CSUS exist for typically developing populations, its use with atypically developing samples is limited: The sample size of the pilot study done with children and adolescents with ASD in North America was not large enough to draw strong conclusions and its age range (10–16 years) was very limited. The study with HI was conducted only with Polish elementary school-aged children (6–12 years). Yet, the other parent-report ToM scales were found to have good psychometric properties among children and adolescents whose ages range between 2 and 18 years. Also, the psychometric properties of the CSUS have been investigated only among preschoolers with typical development in Turkey. The present study examines the psychometric properties of the CSUS with a large sample of children and adolescents with ASD, aged between 8 and 17 years (Study 1) and a sample of children and adolescents with HI using hearing devices, aged between 3 and 12 years (Study 2) in Turkey. Moreover, the psychometric properties of the CSUS have been mostly investigated with its full form (e.g., Białecka-Pikul & Stępień-Nycz, 2019, Brosseau-Liard & Poulin-Dubois, 2018, Jang & Shin, 2018, Smogorzewska et al., 2019). Yet, only the studies with Turkish (Ekerim-Akbulut et al., 2021) and French (Brosseau-Liard & Poulin-Dubois, 2018) children were conducted with the short form. Therefore, this study also aims to demonstrate the use of both short and full forms of the CSUS in ToM assessment.

Study 1

In Study 1, we investigated the psychometric properties of the CSUS-Short form in children and adolescents with ASD. Because false belief understanding tasks have been the most widely used ToM measures in the literature (Bauminger & Kasari, 1999, Joseph & Tager-Flusberg, 2004), we utilized two different versions of false belief understanding tasks to assess the validity of the CSUS-Short form. We used one standard ToM task, the unexpected change false belief understanding task, which has been widely used with typically and atypically developing children and adolescents (e.g., Baron-Cohen et al., 1985, Wimmer & Perner, 1983) and one low-verbal ToM task,

which has previously yielded a reliable pattern of results in other atypically developing samples (e.g., children with HI, Alayli & Yagmurlu, 2014, Woolfe et al., 2002). We also explored the relation of the CSUS-Short form with the children's and adolescents' receptive language and non-verbal intelligence.

Method

Participants

The data were collected from 106 children and adolescents (87 boys) with a diagnosis of ASD, whose ages ranged between 8.00 and 17.11 years ($M_{age} = 12.06$ years, $SD = 2.91$) and their parents (103 mothers, 3 fathers). The ratio of boys to girls was 4.5:1, reflecting the ratio of sex in the ASD population (Christensen et al., 2016). There were no age differences between boys and girls. Children and adolescents who were attending a special education center in Istanbul were recruited for the study. All the participating children and adolescents were diagnosed with ASD by a psychiatrist or neurologist based on DSM-IV-TR criteria (American Psychiatric Association, 2000). The study was conducted with mild and high functioning children and adolescents with ASD who did not have a diagnosis of severe intellectual disability and had a sufficient linguistic capacity to communicate. Demographic information indicated diverse educational backgrounds and employment for parents. Fifty-one mothers (49.0%) and 54 fathers (51.9%) had at least a high school diploma, while 7 mothers (6.7%) and 5 fathers (4.8%) had not graduated from primary school. Seventy-six fathers (71.7%) had a full-time job; while only 23 mothers (21.7%) were employed. The sample reflected the population of parents in Turkey.

Measures

Children's Social Understanding Scale-Short Form (CSUS-Short Form; Tahiroglu et al., 2014) The parents were asked to fill out the Turkish adaptation of the CSUS-Short form (18 items; Ekerim-Akbulut et al., 2021). The items of the CSUS-Short form assess a range of ToM skills and behaviors reflecting children's understanding of mental states (e.g., "When given an undesirable gift, pretends to like it so as not to hurt the other person's feelings."; see Tahiroglu et al., 2014 for a complete list of items). The parents rated each item on a 4-point Likert scale (from 1 = *never* to 4 = *always*; Ekerim-Akbulut et al., 2021). The parents could also select the option 'don't know' in case they had no idea about the behavior asked in the question, which was then coded as missing data. The CSUS-Short form score was computed by averaging the scores across the 18 items, with high scores reflecting better ToM understanding. Reliability

analyses revealed that the CSUS-Short form had a very high internal consistency in this sample ($\alpha = 0.91$).

Theory of Mind Tasks The children's and adolescents' behavioral ToM performance was assessed with one low-verbal and one standard ToM task, measuring their first-order false belief understanding. Although the children and adolescents were expected to have sufficient verbal comprehension skills, the low-verbal ToM task differs from traditional tasks in requiring fewer or no verbal responses. In the low-verbal ToM task (see Woolfe et al., 2002 for details), the participants' understanding of 'reality' vs 'belief' was measured via the use of thought bubbles. These thought bubbles stand as the indications of 'beliefs' of the protagonist rather than reality. To test whether the children and adolescents understood that the 'thought bubble' stood for belief, they were administered one example scenario. In the example scenario, there were two pictures of a boy, one depicting a boy near a dog and the other depicting a boy thinking about a dog by the use of a thought bubble. The children and adolescents were asked to show the picture of the boy thinking about a dog. If they showed the correct picture in the example scenario, they were given two experimental scenarios in which the characters' thoughts differed from reality. For example, in one of the experimental scenarios, the children and adolescents were shown a picture of a boy with a fishing rod. They were told that the child thought he had caught a fish. In the picture, there were also reeds as an obstruction item that covered the caught item. Then, the picture without the obstruction item was shown to the participants, depicting the boy catching a boot. The participants' task was to pick the correct picture depicting the character's belief (what the character was thinking) and reality (the actual situation concealed by the obstruction item). The children and adolescents received a score of 1 for each correct answer in response to experimental scenarios, and their low-verbal ToM scores thus ranged from 0 to 2.

The unexpected location change task (Wimmer & Perner, 1983) was also administered as a standard ToM task to measure children's and adolescents' false belief understanding. This task is frequently used to measure ToM of children and adolescents with ASD. In this task, the participants were told a story about two characters. After one of the characters puts his ball in one location (blue box) and leaves the scene, the other character comes and changes the location of the ball (to the yellow box). After hearing the story, the children and adolescents were asked two control questions about where each character put the ball. The children and adolescents who passed the control questions were asked two experimental questions about where the first character would think the object is and where he would look for the object upon coming back to the scene. Then, the

participants were asked two memory questions about the previous and current locations of the ball. If the participants could not pass the control and/or memory questions, they were excluded from data analysis. If they correctly answered all the control and memory questions, their behavioral ToM score was computed: Children and adolescents received a score of 1 for each correct answer in response to experimental questions, and their ToM scores thus ranged from 0 to 2.

Receptive Language The children's and adolescents' receptive language was measured by the Turkish equivalent of Peabody Picture Vocabulary Test (PPVT), Receptive Language subtest of the Turkish Expressive and Receptive Language Test (TIFALDI-RT; Berument & Güven, 2013). It is a reliable and valid measure for assessing the receptive vocabulary abilities of 2- to 12-year-old typically developing children and adolescents (e.g., Korucu et al., 2016). Similar to the PPVT, the children and adolescents were instructed to select the picture of the target word which was uttered by the experimenter out of four pictures. We slightly changed the standard administration of the test based on the advice from the developers of the test since linguistic capacities of children and adolescents with ASD vary greatly: Rather than starting with the word designated according to chronological age, we started the task with the word designated to one year below the chronological age in this study. For the adolescents who were older than 12 years old, we started with the word designated for 12-years-olds. The children and adolescents received 1 point for each correct answer and the total score was computed by summing up all the points. Because there are no standardized scores available for atypically developing populations, the raw scores were used.

Nonverbal Intelligence The children's and adolescents' nonverbal intelligence was assessed via the Raven Colored Progressive Matrices (Raven, 1938). This tool has been previously used with Turkish children and showed good psychometric properties (e.g., Yagmurlu et al., 2005). The task contained 36 figures with different colored patterns and six possible pieces, only one of which could fill the blank space in the figure and complete the pattern. The children and adolescents were instructed to show the piece completing the figure. The figures of the task were designed in the order of increasing difficulty. The participants were given a score of 1 for each correct answer. The scores were summed to calculate the nonverbal intelligence score, with the values ranging from 0 to 36.

Procedure

Data collection started upon receiving the IRB approval. The participants were recruited through 20 special education

centers in Istanbul, Turkey. The parents who gave a written informed consent were asked to complete the background information form and the CSUS-Short form at home. After the parents completed the forms, the children and adolescents were visited at the special education centers for behavioral assessments. The mean time gap between the parents' completion of the forms and behavioral assessments was 29 days. After receiving a verbal assent from the children and adolescents, the behavioral tasks were presented in the following fixed order: Low-verbal ToM task, standard ToM task, receptive language test, and nonverbal intelligence task. The administration of the tasks lasted for approximately 45 minutes. None of the children and adolescents were excluded from the study due to no cooperation.

Results

Preliminary Analyses of Theory of Mind Measures

In the standard ToM task, 4 participants failed to pass control and/or memory questions; thus, they were excluded from the study (see Table 1 for descriptive statistics).

The CSUS-Short form scores of the children and adolescents ranged from 1.17 to 3.83 with a mean value of 2.32 ($SD = 0.64$). Factor analysis yielded one factor explaining 46% of the variance. This result is consistent with the original scale, in which one factor explained 32% of the variance (Tahiroglu et al., 2014). The CSUS-Short form scores did not differ according to the children's and adolescents' sex and no significant associations were found between the CSUS-Short form and children's and adolescents' age or parents' educational background (see Table 2).

Children and adolescents performed better on the low-verbal ToM task ($M = 1.26$, $SD = 0.81$) than on the standard ToM task ($M = 0.81$, $SD = 0.88$), $F(1, 101) = 21.1$, $p < 0.001$, $\eta^2 = 0.17$. While boys ($M = 1.36$, $SD = 0.77$) performed better on the low-verbal ToM task than girls ($M = 0.83$, $SD = 0.86$), $F(1, 100) = 6.58$, $p = 0.01$, $\eta^2 = 0.06$; there was no sex difference in participants' standard ToM task performance. There were no significant associations between the behavioral ToM tasks and children's and adolescents' age or parents' educational background.

Main Analyses

Due to multiple comparisons (7 comparisons in total), we used Bonferroni correction for statistical adjustment, and the significance level was set at 0.007. Children's and adolescents' performance on low-verbal and standard ToM tasks was positively and significantly correlated (see Table 2). The CSUS-Short form score was positively and significantly associated with the participants' low-verbal ToM performance but not to standard ToM task scores. The

Table 1 Descriptive Statistics for the Measures Used in the ASD Sample (Study 1, $N = 102$)

	<i>M (SD)</i>	<i>Min-Max</i>	<i>n (%)</i>	<i>M_{age} (SD)</i>	<i>M_{CSUS} (SD)</i>
CSUS-Short form	2.32 (0.64)	1.17–3.83			
Low-verbal ToM	1.26 (0.81)	0–2			
0			23 (22.5)	12.0 (2.65)	2.06 (0.50)
1			29 (28.5)	13.5 (2.80)	2.10 (0.62)
2			50 (49)	11.3 (2.89)	2.57 (0.63)
Standard ToM	0.81 (0.88)	0–2			
0			50 (49)	12.2 (2.73)	2.22 (0.64)
1			21 (20.6)	11.5 (3.25)	2.37 (0.57)
2			31 (30.4)	12.4 (3.05)	2.44 (0.68)
Receptive language	72.4 (26.4)	0–102			
Nonverbal intelligence	20.2 (10.3)	0–36			
Child's age	12.1 (2.93)	8.00–17.11			
Child's sex					
Girl			18 (17.6)	12.3 (2.94)	2.28 (0.59)
Boy			84 (82.4)	12.1 (2.94)	2.33 (0.66)

CSUS Children's Social Understanding Scale, ToM Theory of mind

Table 2 Zero-Order Correlations after Bonferroni Correction in the ASD Sample (Study 1, $N = 102$)

	CSUS-Short form	1	2	3	4	5	6
1. Low-verbal ToM	0.36**	–					
2. Standard ToM	0.15	0.31**	–				
3. Receptive language	0.40**	0.48**	0.33*	–			
4. Nonverbal intelligence	0.29*	0.39**	0.23	0.69**	–		
5. Child's age	–0.12	–0.17	0.01	–0.16	–0.01	–	
6. Maternal education	0.07	–0.04	0.07	0.24*	0.17	0.02	–
7. Paternal education	–0.09	0.01	–0.03	0.09	0.02	–0.02	0.66**

The bold values represent the significant associations.

CSUS Children's Social Understanding Scale, ToM Theory of mind

* $p < 0.007$, ** $p < 0.001$

CSUS-Short form score and ToM performance on the low-verbal and standard ToM tasks were all positively associated with the children's and adolescents' receptive language. Also, the CSUS-Short form score and low-verbal ToM performance were positively associated with the children's and adolescents' nonverbal intelligence. The association between the CSUS-Short form score and low-verbal ToM performance remained significant after controlling for the children's and adolescents' receptive language, $r(102) = 0.20$, $p = 0.04$, and nonverbal intelligence, $r(102) = 0.28$, $p = 0.005$.

Discussion

In Study 1, the CSUS-Short Form showed a very high internal consistency. We measured children's and adolescents' ToM performance, receptive language, and nonverbal intelligence to evaluate the validity of the scale. As would be expected based on the literature (e.g., Kerr & Durkin,

2004), our findings showed that the children and adolescents with ASD performed better on the low-verbal ToM task than on the standard ToM task. These results complement the existing literature that ToM performance of individuals with ASD varies according to task demands (Hutchins et al., 2016, Van Herwegen et al., 2013). The children's and adolescents' CSUS-Short form score was also positively associated with their performance on the low-verbal ToM task, but not with performance on the standard ToM task. Considering that the low-verbal ToM task may more accurately capture individual differences in ToM abilities of children and adolescents with ASD than the standard ToM task, our results provide important information regarding the validity of the CSUS-Short form in this sample.

We found that the CSUS-Short form score and ToM performance were positively associated with receptive language and nonverbal intelligence, supporting the link of ToM with verbal abilities (Frith et al., 1994, Happé, 1995,

Milligan et al., 2007, Wellman, 2014, Yirmiya et al., 1998) and intelligence level (Happé, 1995) in both typical and atypical samples. We also found that neither the CSUS-Short form scores nor ToM performance were significantly correlated with the participants' chronological age. This is consistent with the literature showing that ToM ability is more closely linked to language and intelligence than to chronological age in ASD (Happé, 1995, Yirmiya et al., 1998). To eliminate the potential influence of the children's and adolescents' linguistic and nonverbal cognitive abilities on parents' evaluation of their children's ToM, we examined the association between the CSUS-Short form score and low-verbal ToM performance via controlling for receptive language and nonverbal intelligence, separately. We found that the associations continued to be significant after controlling for the participants' receptive language and nonverbal intelligence. This indicates that the parents' assessment of their children's ToM on the CSUS-Short form is, to a degree, independent of their linguistic and intellectual abilities. These results demonstrate the validity of the CSUS-Short form in this sample.

Study 2

In Study 2, we investigated the psychometric properties of the CSUS-Full form in assessing ToM of children and adolescents with hearing impairment (HI). We examined the internal consistency and validity of the scale through its correlations with performance on behavioral ToM tasks: Two low-verbal ToM tasks and the ToM scale consisting of five ToM tasks. The children and adolescents with hearing impairment all used spoken language as their daily mode of communication. Thus, instead of using non-verbal ToM tasks, we used low-verbal ToM tasks to minimize the verbal load of the tasks while still acknowledging the importance of verbal communication. We also explored the relation of the CSUS-Full form with the participants' receptive language scores.

Method

Participants

We collected data from 70 children and adolescents (33 boys) with HI whose ages ranged between 3 and 12 years ($M_{age} = 6.8$ years, $SD = 2.32$) and their mothers. Children's and adolescents' degree of hearing loss ranged between mild (hearing loss from around 26 to 40 decibels) and profound (hearing loss more than 91 decibels). The devices they used were cochlear implants ($N = 45$) or hearing aids ($N = 25$), with the mean age of starting device use being 3.05 years ($SD = 1.85$). All children and adolescents used

spoken language as their main mode of communication with little or no sign language knowledge.

Recruitment of the participants was held in four special education and rehabilitation centers, in three big cities (Istanbul, Eskisehir, and Kahramanmaras), located in different regions of Turkey. Each center's education medium was spoken Turkish language. Based on parent reports, all participants had no other disability than HI. The parents of all children and adolescents (but 1) were hearing individuals. The majority of the mothers were stay-at-home parents (87%) and had education at the primary school level (43.5%). On the other hand, the majority of the fathers had a paid job (94%) and had education at the high school level (35%). Other than mothers' education level, the sample reflected the population of parents in Turkey.

Measures

Children's Social Understanding Scale-Full Form (CSUS-Full form; Tahiroglu et al., 2014) Turkish adaptation of the full form of the CSUS (42 items; Tahiroglu & Yagmurlu, 2014) was administered to the parents. The application of the full form is the same as the short form used in Study 1. Reliability analyses revealed that the CSUS-Full form had a very high internal consistency in this sample ($\alpha = 0.92$).

Theory of Mind Tasks The children's and adolescents' behavioral ToM performance was assessed via two low-verbal tasks measuring false belief understanding and the ToM scale with five tasks created by Wellman and Liu (2004). As low-verbal ToM tasks, we used the unexpected contents and thought bubble tasks, both including stories with pictures that required little or no verbal prompting to reduce the cognitive load. In the unexpected contents task (de Villiers & de Villiers, 2012), the children and adolescents were shown picture sequences depicting one of the characters substitute unusual objects (pencils) with objects one would expect to find in a familiar container (candy box). Then, the participants were asked to tell what the other character would think is inside the container, candy or pencils. The participants received 1 point for the correct answer (candy). In the thought bubble task, the extended version of the low-verbal ToM task described in Study 1 (Woolfe et al., 2002) was used. In this task, in addition to two false-belief scenarios, two true-belief scenarios, in which there was no conflict with reality, were administered and all four scenarios were given in a counterbalanced order. We calculated true belief and false belief scores separately, and since the focus of this study was false belief understanding, we only used false belief scores in the analysis. To pass each trial and get a score of 1, the children and adolescents needed to pick the correct picture depicting the character's belief and reality. As in Study 1, the

participants' scores for each false-belief scenario were aggregated to obtain a composite score of behavioral ToM, ranging from 0 to 2 points. To calculate the low-verbal ToM score, the scores of the unexpected contents and low-verbal ToM tasks were aggregated, resulting in a score ranging between 0 and 3.

We also administered the ToM scale created by Wellman and Liu (2004), which consisted of five tasks: diverse desires, diverse beliefs, knowledge access, content false belief, and hidden emotion (see Wellman & Liu, 2004, for details). In the diverse desires task, the experimenter showed the participants two food items (a carrot and a cookie) and asked which one the participant favors the most. Depending on the choice, the experimenter told the participant that another child liked the other food item. Then, the experimenter explained that it was snack time and asked which food item the other child would select. If the participant chose the other child's favorite food item, then the participant passed the task, indicating that they distinguished their own desire from those of others'. Similarly, in the diverse beliefs task, the participant and another protagonist held different beliefs on the same situation (e.g. the hiding place of a cat). Knowledge access task required the participants to differentiate their own and others' knowledge about the content of a box. For the contents false belief task, participants were judged whether another person had a false belief about the content of a pencil box. Lastly, in the hidden emotion task, the children and adolescent were judged whether the emotion displayed by an individual was different from that person's actual emotion. Wellman and Liu's (2004) ToM scale was previously used with Turkish children and showed good psychometric properties (e.g., Etel & Yagmurlu, 2015, Korucu et al., 2016). For each item, a control question and a target question were asked to the participants. Both questions had to be correctly answered to get a score of 1. To obtain ToM scale composite score, scores of each task were aggregated, ranging from 0 to 5.

Receptive Language The participants' receptive language was measured by the same receptive vocabulary task used in Study 1. We followed the standard procedure and the task started with the word designated according to the chronological age of the participants. We used the same scoring procedure explained in Study 1 and used the raw scores for analyses.

Procedure

Data collection started upon receiving the IRB approval. After receiving the informed consent from the parents and directors of the special education centers, the CSUS-Full form was given to the parents through the teachers in the

special education centers and collected within 2 weeks. The mothers completed the scale by themselves, yet if needed, clarifications were provided by the experimenter. For two illiterate mothers, the questions were read out loud face-to-face/on the phone and oral responses were noted down on the form by the experimenter. The children and adolescents had limited sign language skills and were able to hear and follow all of the instructions. The participants also had a practice with spoken language which was the medium of their everyday communication and education in the rehabilitation centers. Thus, after receiving an assent from the children and adolescents, the tasks were administered in spoken language by a hearing experimenter who was competent in sign language. The order of the tasks was as follows: Receptive language test, low-verbal ToM tasks (the unexpected contents and false belief understanding tasks, respectively), and ToM scale. Before administering the ToM scale, a 5-minute break was given. None of the children and adolescents were excluded from the study due to no cooperation.

Results

Preliminary Analyses of Theory of Mind Measures

All children and adolescents passed control and/or memory questions in the behavioral ToM tasks (see Table 3 for descriptive statistics).

The CSUS-Full form scores of the children and adolescents with HI ranged between 1.69 and 3.83 ($M = 2.81$, $SD = 0.54$). Factor analysis yielded one factor accounting for 31% of the variance, which is consistent with the original scale's one-factor solution (Tahiroglu et al., 2014). There were no sex differences in CSUS-Full form scores. Zero-order correlations revealed that children's and adolescents' age was positively correlated with the CSUS-Full form (see Table 4). The CSUS-Full form was not significantly correlated with the parents' educational background.

The participants' performance on low-verbal ToM tasks and ToM scale did not differ according to their sex. The participants' age was positively associated with their low-verbal ToM task performance, but not with the ToM scale (see Table 4). The parents' educational background was not significantly associated with low-verbal ToM but was positively correlated with the ToM scale.

Main Analyses

Due to multiple comparisons (7 comparisons in total), Bonferroni correction was used and the significance level was adjusted to 0.007. The CSUS-Full form score was positively correlated with the participants' performance on low-verbal ToM, ToM scale, and receptive language (see

Table 3 Descriptive Statistics for the Measures Used in the HI Sample (Study 2, $N = 70$)

	<i>M (SD)</i>	<i>Min-Max</i>	<i>n (%)</i>	<i>M_{age} (SD)</i>	<i>M_{CSUS} (SD)</i>
CSUS-Full form	2.81 (0.54)	1.69–3.83			
Low-verbal ToM	1.70 (0.99)	0–3			
0			7 (10)	6.1 (2.03)	2.46 (0.40)
1			27 (38.6)	5.8 (2.04)	2.69 (0.54)
2			16 (22.9)	7.56 (2.36)	2.88 (0.53)
3			20 (28.6)	7.6 (2.36)	3.02 (0.53)
ToM scale	2.37 (1.36)	0–5			
0			6 (8.6)	6.5 (2.42)	2.53 (0.31)
1			13 (18.6)	6.69 (2.68)	2.62 (0.52)
2			21 (30)	6.23 (1.94)	2.83 (0.54)
3			13 (18.6)	7.23 (2.54)	2.87 (0.61)
4			13 (18.6)	7 (2.27)	3.05 (0.41)
5			4 (5.7)	8.5 (2.88)	3.16 (0.49)
Receptive language	50.8 (21.8)	14–97			
Degree of hearing loss	168.1 (34.1)	36–146			
Child’s age	6.8 (2.32)	3.00–12.00			
Child’s sex					
Girl			37 (52.9)	6.9 (2.34)	2.8 (0.48)
Boy			33 (47.1)	6.6 (2.32)	2.79 (0.62)

CSUS Children’s Social Understanding Scale, ToM Theory of mind

Table 4 Zero-Order Correlations after Bonferroni Correction in the HI Sample (Study 2, $N = 70$)

	CSUS-Full form	1	2	3	4	5	6
1. Low-verbal ToM	0.32*	–					
2. ToM scale	0.39**	0.23	–				
3. Receptive language	0.49**	0.25	0.38*	–			
4. Child’s age	0.35*	0.30	0.16	0.43**	–		
5. Degree of hearing loss	–0.25	–0.28	–0.40*	–0.44*	–0.11	–	
6. Maternal education	0.13	0.06	0.28	0.12	–0.10	–0.19	–
7. Paternal education	0.19	0.04	0.23	0.16	–0.05	0.52**	–0.18

The bold values represent the significant associations.

CSUS Children’s Social Understanding Scale, ToM Theory of mind

* $p < 0.007$, ** $p < 0.001$

Table 4). In addition, performance on ToM scale was positively associated with receptive language. The association between the CSUS-Full form score and ToM performance continued to be significant after controlling for the participants’ age, $r(70) = 0.26$, $p = 0.035$ for low-verbal ToM and $r(70) = 0.36$, $p = 0.002$ for ToM scale. Also, the association was significant after controlling for the participants’ receptive language, $r(70) = 0.25$, $p = 0.047$ for low-verbal ToM and $r(70) = 0.25$, $p = 0.037$ ToM scale.

Discussion

In Study 2, the CSUS-Full form showed a high internal consistency. To evaluate the validity of the CSUS-Full form, we measured children’s and adolescents’ ToM

performance via both low-verbal and standard ToM tasks and assessed their receptive language. The results demonstrated that participants’ CSUS-Full form scores were associated with their performance on low-verbal and standard ToM tasks. The CSUS-Full form was also positively associated with the participants’ age and receptive language, supporting the link between ToM with chronological age (Peterson & Wellman, 2009) and linguistic abilities (de Villiers & de Villiers, 2012, Woolfe et al., 2002) in samples with typical development and HI. The associations between the CSUS-Full form score and ToM performance continued to be significant when we controlled for the children’s and adolescents’ age and receptive language, separately. Similar to Study 1, this finding suggests that the parents’ assessment of their children’s ToM on the CSUS-Full form is, to a

degree, independent of their age or linguistic abilities. These results demonstrate the validity of the CSUS-Full form in this sample.

Despite the association between the CSUS-Full form and performance on the low-verbal and ToM scale, there was no significant correlation between low-verbal ToM and standard ToM tasks. One possible explanation could be that these tasks measured different types of ToM. While low-verbal ToM tasks focused on measuring false belief understanding, the ToM scale measured additional aspects of ToM, which could give rise to a lack of association between the two measures. To test this idea, we investigated the association between low-verbal ToM tasks and false belief understanding performance from the ToM scale. Yet, there was no correlation between them. The lack of a correlation also indicates that it might be the task demands that cause a non-significant association between the two measures (Hutchins et al., 2017). Whereas low-verbal ToM tasks minimize the verbal load of the task, standard ToM tasks have longer scenarios and questions. This difference in the verbal load requires participants to pay more attention to understand, process, and memorize the scenarios and questions. The difficulty level of the tasks in the ToM scale increased, respectively, to represent developmental course of ToM (Wellman & Liu, 2004). The last task in the ToM scale, which was the hidden emotion, was given after the false belief understanding task and had the highest level of verbal load and difficulty in comparison to the other tasks in the ToM scale (Wellman & Liu, 2004). Thus, to assess whether task demands led to insignificant correlations between the two measures, we reran the analyses excluding the hidden emotion task. Our findings showed a significant correlation with low-verbal ToM ($r = 0.24$, $p = 0.047$). These findings demonstrate the importance of task demands on children's and adolescents' ToM performance.

General Discussion

The current research aimed to examine psychometric properties of the CSUS in two different atypically developing populations with wider age ranges than the ones tested in the previous studies. We conducted two separate studies with Turkish children and adolescents with autism spectrum disorder (ASD) and hearing impairment (HI) using hearing devices. The factor analyses revealed a one-factor structure of the CSUS in both studies, as was the case in the original scale (Tahiroglu et al., 2014), Turkish version of the CSUS-Short form used among typically developing children (Ekerim-Akbulut et al., 2021), and adaptation of the CSUS to the other cultures (e.g., Smogorzewska et al., 2019). Although the CSUS measures different aspects of ToM, these findings support the recommendation of

Tahiroglu et al. (2014) for the use of a total score rather than the subscale scores. This can be interpreted as the possibility of the parents' evaluation of their children's ToM based on their simultaneous use of different aspects in everyday interactions.

Our findings also demonstrated high internal consistencies for both short and full forms of the scale, similar to the previous studies conducted with the CSUS (e.g., Białecka-Pikul & Stępień-Nycz, 2019, Brosseau-Liard & Poulin-Dubois, 2018, Ekerim-Akbulut et al., 2021, Smogorzewska et al., 2019, Tahiroglu et al., 2014). These results revealed that the CSUS was able to reliably assess individual differences in ToM in these atypically developing samples (as it does in typically developing preschool-age samples).

As far as the validity was concerned, the short and full forms of the CSUS were correlated with performance on behavioral ToM tasks in line with the literature. Consistent with the earlier studies on the psychometric properties of the CSUS (e.g., Białecka-Pikul & Stępień-Nycz, 2019, Brosseau-Liard & Poulin-Dubois, 2018, Smogorzewska et al., 2019, Tahiroglu et al., 2014), including its Turkish adaptation (Ekerim-Akbulut et al., 2021), we found moderate and positive correlations between the CSUS and ToM performance, specifically between the CSUS-Short form and low-verbal ToM in the ASD sample and between the CSUS-Full form and ToM performance (on both low-verbal and standard tasks) in the HI sample. Correlation analyses also showed significant associations of the CSUS-Short form with receptive language and nonverbal intelligence in the ASD sample and of the CSUS-Full form with age and receptive language in the HI sample. Similar associations have been also shown in the extant literature, namely positive associations of ToM with nonverbal intelligence in ASD sample (Happé, 1995), with age in HI sample (Peterson & Siegal, 2000), and with receptive language in atypically developing samples (Milligan et al., 2007, Wellman, 2014). Partial correlation analyses revealed that the association between the CSUS and ToM performance remained significant. These findings demonstrate that the CSUS measures children's and adolescents' ToM controlling for their age, nonverbal intelligence, and receptive language, supporting the validity of the short and full forms of the CSUS in these samples. Thus, the results highlight the use of the CSUS for the assessment of children's and adolescents' mental state understanding.

One of the strengths of this study was the use of both low-verbal and standard ToM tasks. Therefore, in atypically developing samples, we investigated differential associations of ToM assessment tools. In the ASD sample, the participants' performance on low-verbal and standard ToM tasks were significantly associated, which might be due to the fact that both types of tasks focused on the

children's and adolescents' false belief understanding. On the other hand, we found that the CSUS-Short form was associated with the participants' performance on low-verbal ToM, but not on standard ToM, tasks. The differential association may be due to task demands (Kaland et al., 2008, Tager-Flusberg, 2007). Similar to the low-verbal ToM task, the CSUS emphasizes children's and adolescents' ToM more than their linguistic and cognitive abilities. Yet, the standard ToM tasks typically demand children and adolescents to use their other abilities to understand, process, and respond to the questions, which may have caused the non-significant association with the CSUS-Short form. Moreover, in the sample with HI, children's and adolescents' performance on low-verbal and ToM scale were not significantly correlated. This may be due to the different characteristics of the tasks like their focus on ToM aspects or linguistic demands. To test the idea, we investigated the relation of the low-verbal ToM tasks with only false belief understanding task in the ToM scale and the tasks without the highest level of verbal load. When we excluded the most verbally-loaded task, the hidden emotion, and rerun the analyses, we found a significant association between low-verbal and ToM scale, revealing the importance of task demands. Moreover, in this sample, the CSUS-Full form was associated with the participants' performance on low-verbal and ToM scale, supporting the notion that all of these measurements assess ToM, albeit from different perspectives. To summarize, though these tools measure ToM, the evaluation of children's and adolescents' ToM might vary according to task demands like general verbal and cognitive loads of a task (Hutchins et al., 2016, 2017, Van Herwegen et al., 2013).

The other strengths of this study were the wide age range and sample size. This study expanded the age range of the scale and revealed that the Turkish version of the CSUS could be used for children and adolescents with atypical development whose ages ranged from 3 to 18 years. Yet, in the ASD and HI studies, we did not have a control group to compare typical and atypical development. If we had a control group, we would expect that children and adolescents with typical development would have higher CSUS scores than children and adolescents with atypical development. Thus, in future studies, comparison group(s) (matched on chronological age, linguistic abilities, and/or intelligence) such as children and adolescents with typical development, intellectual disability, communication disorder, and Down's syndrome may be included. It would have been interesting to make comparisons with the scores of children and adolescents with typical development from previous studies. However, no age-matched Turkish children and adolescents have been tested so far. Thus, in the future, it would be interesting to see if the CSUS would have discriminant validity.

One of the aims of the study was to highlight the use of short and full forms of the CSUS in atypical populations. This study demonstrated the reliability and validity of both forms in ASD and HI samples. Yet, the design of the studies prevented the comparison between the different forms of the CSUS and behavioral ToM tasks for the atypical samples. The ASD and HI studies were conducted at different time points with different research questions. The ASD study was an extensive study in which we administered a series of questionnaires to the parents; thus, we preferred to use the short form of the CSUS to decrease the burden on the parents. On the other hand, in the HI study, we gave only one questionnaire, the CSUS-Full form, to the parents and aimed to measure different aspects of the ToM. To provide an opportunity for comparison, in the HI study, we picked the questions of the short form and calculated the CSUS-Short form score. Our findings showed that the CSUS-Short form was significantly associated with low-verbal ToM ($r = 0.31$, $p < 0.05$) and ToM scale ($r = 0.39$, $p = 0.001$). The Fisher's z test also demonstrated that these correlation coefficients were not significantly different from the ones we found with the full form ($z = 0.06$, $p = 0.47$ for low-verbal ToM, $z = 0.00$, $p = 0.50$ for ToM scale). These results demonstrate that there is a significant association between the CSUS and behavioral ToM tasks regardless of the length of the forms.

Our findings showed that the CSUS was more strongly associated with receptive language than behavioral ToM measures in the atypically developing samples. Moreover, receptive language was associated with all of the ToM tools, with the exception of low-verbal ToM in Study 2. These findings suggest that children's and adolescents' linguistic abilities might be more prominent for assessing ToM. While discussing the limitations of the behavioral ToM assessments, the importance of children's and adolescents' linguistic abilities was particularly highlighted. Children with atypical development mostly rely on their linguistic abilities more so than their social insights since they have broader difficulties in understanding social interactions (Tager-Flusberg, 2007). This pattern might not be specific to laboratory studies but can be observed in their everyday interaction. Also, parents might perceive their children who understand more words as more competent at thinking about their and others' mental states (Białecka-Pikul & Stępień-Nycz, 2019). Therefore, children's and adolescents' linguistic abilities might influence not only their performance on behavioral ToM tasks but also parents' perception, as well. These might also explain why we found significant but moderate associations between the CSUS and ToM performance. To better understand the role of language, future studies may investigate the effect of linguistic abilities on ToM. Additionally, the use of receptive language task as the only indicator of children's and

adolescents' linguistic abilities was another limitation. For instance, research shows that ToM can be related to other aspects of language development, such as syntactic development and conversational experience (Colle et al., 2007, de Villiers & de Villiers, 2012, Kelley et al., 2006). Given the known difficulties in children and adolescents with ASD and HI in responding to questions, sharing and requesting information (e.g., Hutchins et al., 2017, Tager-Flusberg, 1996, Woolfe et al., 2002), and producing narratives (e.g., Kelley et al., 2006), we recommend future studies to investigate the relationship between measures of ToM (especially the CSUS) and different aspects of language, such as explicit language focusing on syntax or conversational opportunities.

The CSUS offers a promising alternative for special educators, psychotherapists, and researchers to practically assess ToM of children and adolescents with atypical development. As indicated by its correlations with low-verbal and standard ToM tasks, the CSUS assesses ToM from a broader perspective, alleviating the need to administer multiple tasks, which might potentially increase verbal and cognitive load. Moreover, the CSUS can be advantageous for longitudinal and intervention studies when investigating the effect of intervention and change in ToM of children and adolescents with atypical development between sessions. Lastly, in contrast to behavioral ToM tasks that depend on relatively advanced cognitive capacities, a parent-report measurement such as the CSUS can be especially useful for children and adolescents with atypical development who have little or no verbal ability and/or severe cognitive deficits (Colle et al., 2007) and severe or profound intellectual disability (Boat & Wu, 2015). Future studies may investigate the psychometric properties of the CSUS among these children and adolescents.

To conclude, our study extended the literature by demonstrating that the CSUS can be used for atypically developing populations, specifically children and adolescents with ASD and HI. We hope that further validation of the CSUS in other atypically developing populations, such as intellectual disability and specific language impairment, wider age ranges, and different cultures will give rise to a deeper understanding of how social-communicative and ToM abilities develop in children and adolescents, and widen our understanding of how universal this ability is as well as how it is shaped by the environment.

Author Contributions B.B., A.A., and B.S. contributed to study conception and design. B.B. and A.A. prepared the materials and collected data. Data analyses were performed by all authors. The first draft of the manuscript was written by B.B. and A.A. and all authors commented on previous versions of the manuscript. All authors read and approved the final manuscript.

Compliance with Ethical Standards

Conflict of Interest The authors declare no competing interests.

Ethical Approval The studies were approved by the institutional research ethics committee of Koç University. They were conducted in accordance with the ethical standards of the American Psychological Association and with the 1964 Helsinki declaration and its later amendments or comparable ethical standards.

Consent to Participate Written informed consent was obtained from the parents and assent was obtained from children prior to the behavioral assessments.

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

References

- Ahmadi, S. Z. Z., Jalaie, S., & Ashayeri, H. (2015). Validity and reliability of published comprehensive theory of mind tests for normal preschool children: A systematic review. *Iranian Journal of Psychiatry, 10*, 214–224.
- Alayli, A., & Yagmurlu, B. (2014). *Sequential acquisition of theory of mind in Turkish deaf children* [Poster presentation]. 18th National Congress of Psychology, Bursa, Turkey.
- American Psychiatric Association (2000). *Diagnostic and statistical manual of mental disorders* (4th ed. text rev.). American Psychiatric Publishing.
- Astington, J. W. (2001). The future of theory-of-mind research: Understanding motivational states, the role of language, and real-world consequences. *Child Development, 72*, 685–687. <https://doi.org/10.1111/1467-8624.00305>.
- Astington, J. W. (2003). Sometimes necessary, never sufficient: False-belief understanding and social competence. In B. Repacholi & V. Slaughter (Eds.), *Individual differences in theory of mind: Implications for typical and atypical development* (pp. 13–38). Psychology Press.
- Astington, J. W., & Jenkins, J. M. (1999). A longitudinal study of the relation between language and theory-of-mind development. *Developmental Psychology, 35*, 1311–1320. <https://doi.org/10.1037/0012-1649.35.5.1311>.
- Baillargeon, R., Scott, R. M., & He, Z. (2010). False-belief understanding in infants. *Trends in Cognitive Sciences, 14*, 110–118. <https://doi.org/10.1016/j.tics.2009.12>.
- Baron-Cohen, S., Jolliffe, T., Mortimore, C., & Robertson, M. (1997). Another advanced test of theory of mind: Evidence from very high functioning adults with autism or Asperger syndrome. *Journal of Child Psychology and Psychiatry, 38*, 813–822. <https://doi.org/10.1111/j.1469-7610.1997.tb01599.x>.
- Baron-Cohen, S., Leslie, A. M., & Frith, U. (1985). Does the autistic child have a “theory of mind”? *Cognition, 21*, 37–46. [https://doi.org/10.1016/0010-0277\(85\)90022-8](https://doi.org/10.1016/0010-0277(85)90022-8).
- Bauminger, N., & Kasari, C. (1999). Brief report: Theory of mind in high-functioning children with autism. *Journal of Autism and Developmental Disorders, 29*, 81–86. <https://doi.org/10.1023/A:1025974701090>.
- Begeer, S., Howlin, P., Hoddenbach, E., Clauser, C., Lindauer, R., Clifford, P., Gevers, C., Boer, F., & Koot, H. M. (2015). Effects and moderators of a short theory of mind intervention for children with autism spectrum disorder: A randomized controlled trial. *Autism Research, 8*, 738–748. <https://doi.org/10.1002/aur.1489>.

- Begeer, S., Malle, B. F., Nieuwland, M. S., & Keysar, B. (2010). Using theory of mind to represent and take part in social interactions: Comparing individuals with high-functioning autism and typically developing controls. *European Journal of Developmental Psychology*, 7, 104–122. <https://doi.org/10.1080/17405620903024263>.
- Begeer, S., Rieffe, C., Meerum-Terwogt, M., & Stockmann, L. (2003). Theory of mind-based action in children from the autism spectrum. *Journal of Autism and Developmental Disorders*, 33, 479–487. <https://doi.org/10.1023/A:1025875311062>.
- Berument, S. K., & Güven, A. G. (2013). Turkish Expressive and Receptive Language Test: I. standardization, reliability and validity study of the receptive vocabulary sub-scale. *Turkish Journal of Psychiatry*, 24, 192–201.
- Białecka-Pikul, M., & Stępień-Nycz, M. (2019). Parent-report measure of theory of mind: Adaptation to Polish of the Children's Social Understanding Scale. *European Journal of Developmental Psychology*, 16, 318–326. <https://doi.org/10.1080/17405629.2017.1406850>.
- 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, 1907–1930. <https://doi.org/10.1007/s10803-008-0585-3>.
- Boat, T., & Wu, J. (2015). Clinical characteristics of intellectual disabilities. In *Mental disorders and disabilities among low-income children* (pp. 169–178). The National Academy of Sciences.
- Brent, E., Rios, P., Happé, F., & Charman, T. (2004). Performance of children with autism spectrum disorder on advanced theory of mind tasks. *Autism*, 8, 283–299. <https://doi.org/10.1177/1362361304045217>.
- Brousseau-Liard, P., & Poulin-Dubois, D. (2018). Fiabilité et validité de l'Échelle de compréhension sociale des enfants [Reliability and validity of the French adaptation of the Children's Social Understanding Scale]. *Psychologie Française*. <https://doi.org/10.1016/j.psfr.2018.01.003>.
- Christensen, D. L., Braun, K. V. N., Baio, J., Bilder, D., Charles, J., Constantino, J. N., Daniels, J., Durkin, M. S., Fitzgerald, R. T., Kurzios-Spencer, M., Lee, L., Pettygrove, S., Robinson, C., Schulz, E., Wells, C., Wingate, M. S., Zahorodny, W., & Yeargin-Allsopp, M. (2016). Prevalence and characteristics of Autism Spectrum Disorder among children aged 8 years-Autism and Developmental Disabilities Monitoring Network, 11 sites, United States, 2012. *Morbidity and Mortality Weekly Report (MMWR)*, 65, 1–23. <https://doi.org/10.15585/mmwr.ss6503a1>.
- Clemmensen, L., Jepsen, J. R. M., van Os, J., Blijd-Hoogewys, E. M. A., Rimvall, M. K., Olsen, E. M., Rask, C. U., Bartels-Velthuis, A. A., Skovgaard, A. M., & Jeppesen, P. (2018). Are theory of mind and bullying separately associated with later academic performance among preadolescents?. *British Journal of Educational Psychology*. <https://doi.org/10.1111/bjep.12263>.
- Colle, L., Baron-Cohen, S., & Hill, J. (2007). Do children with autism have a theory of mind? A non-verbal test of autism vs. specific language impairment. *Journal of Autism and Developmental Disorders*, 37, 716–723. <https://doi.org/10.1007/s10803-006-0198-7>.
- de Villiers, J. G., & de Villiers, P. A. (2000). Linguistic determinism and the understanding of false beliefs. In P. Mitchell & K. J. Riggs (Eds.), *Children's reasoning and the mind* (pp. 191–228). Psychology Press.
- de Villiers, P., & de Villiers, J. G. (2012). Deception dissociates from false belief reasoning in deaf children: Implications for the implicit versus explicit theory of mind distinction. *British Journal of Developmental Psychology*, 30, 188–209. <https://doi.org/10.1111/j.2044-835X.2011.02072.x>.
- Ekerim-Akbulut, M., Etel, E., Tahiroğlu, D., & Selçuk, B. (2021). Children's Social Understanding Scale Short-form: Adaptation to Turkish sample. *Early Education and Development*. <https://doi.org/10.1080/10409289.2021.2013668>.
- Etel, E., & Yagmurlu, B. (2015). Social competence, theory of mind, and executive function in institution-reared Turkish children. *International Journal of Behavioral Development*, 39, 519–529. <https://doi.org/10.1177/0165025414556095>.
- Frick, P. J., Barry, C. T., & Kamphaus, R. W. (2010). *Clinical assessment of child and adolescent personality and behavior*. Springer.
- Frith, U., Happé, F., & Siddons, F. (1994). Autism and theory of mind in everyday life. *Social Development*, 3, 108–123. <https://doi.org/10.1111/j.1467-9507.1994.tb00031.x>.
- Gioia, G., Isquith, P., Guy, S., & Kenworthy, L. (2000). *BRIEF-Behavior Rating Inventory of Executive Function, Professional manual*. Psychological Assessment Resources.
- Happé, F., & Frith, U. (1996). The neuropsychology of autism. *Brain*, 119, 1377–1400. <https://doi.org/10.1093/brain/119.4.1377>.
- Happé, F. G. E. (1995). The role of age and verbal ability in the theory of mind task performance of subjects with autism. *Child Development*, 66, 843–855. <https://doi.org/10.2307/1131954>.
- Hoogenhout, M., & Malcolm-Smith, S. (2014). Theory of mind in autism spectrum disorder: Does DSM classification predict development? *Research in Autism Spectrum Disorders*, 8, 597–607. <https://doi.org/10.1016/j.rasd.2014.02.005>.
- Hutchins, T. L., Allen, L., & Schefer, M. (2017). Using the theory of mind inventory to detect a broad range of theory of mind challenges in children with hearing loss: A pilot study. *Deafness & Education International*, 19, 2–12. <https://doi.org/10.1080/14643154.2016.1274089>.
- Hutchins, T. L., Prelock, P. A., & Bonazinga, L. (2012). Psychometric evaluation of the Theory of Mind Inventory (ToMI): A study of typically developing children and children with autism spectrum disorder. *Journal of Autism and Developmental Disorders*, 42, 327–341. <https://doi.org/10.1007/s10803-011-1244-7>.
- Hutchins, T. L., Prelock, P. A., Morris, H., Benner, J., LaVigne, T., & Hoza, B. (2016). Explicit vs. applied theory of mind competence: A comparison of typically developing males, males with ASD, and males with ADHD. *Research in Autism Spectrum Disorders*, 21, 94–108. <https://doi.org/10.1016/j.rasd.2015.10.004>.
- Imuta, K., Henry, J. D., Slaughter, V., Selçuk, B., & Ruffman, T. (2016). Theory of mind and prosocial behavior in childhood: A meta-analytic review. *Developmental Psychology*, 52, 1192–1205. <https://doi.org/10.1037/dev0000140>.
- Jang, M., & Shin, N. (2018). Effects of maternal parenting behavior and preschoolers' executive function on preschoolers' theory of mind. *Korean Journal of Child Studies*, 39, 103–117. <https://doi.org/10.5723/kjcs.2018.39.1.103>.
- Joseph, R. M., & Tager-Flusberg, H. (2004). The relationship of theory of mind and executive functions to symptom type and severity in children with autism. *Development and Psychopathology*, 16, 137–155. <https://doi.org/10.1017/s095457940404444x>.
- Kaland, N., Callesen, K., Møller-Nielsen, A., Mortensen, E. L., & Smith, L. (2008). Performance of children and adolescents with Asperger syndrome or high-functioning autism on advanced theory of mind tasks. *Journal of Autism and Developmental Disorders*, 38, 1112–1123. <https://doi.org/10.1007/s10803-007-0496-8>.
- Kelley, E., Paul, J., Fein, D., & Naigles, L. (2006). Residual language deficits in children with a history of autism. *Journal of Autism and Developmental Disorders*, 36, 807–828. <https://doi.org/10.1007/s10803-006-0111-4>.
- Kerr, S., & Durkin, K. (2004). Understanding of thought bubbles as mental representations in children with autism: Implications for

- theory of mind. *Journal of Autism and Developmental Disorders*, 34, 637–648. <https://doi.org/10.1007/s10803-004-5285-z>.
- Kleinman, J., Marciano, P. L., & Ault, R. L. (2001). Advanced theory of mind in high-functioning adults with autism. *Journal of Autism and Developmental Disorders*, 31, 29–36. <https://doi.org/10.1023/A:1005657512379>.
- Korucu, I., Selcuk, B., & Harma, M. (2016). Self-regulation: Relations with theory of mind and social behaviour. *Infant and Child Development*, 26, 1–23. <https://doi.org/10.1002/icd.1988>.
- Kovács, A. M., Téglás, E., & Endress, A. D. (2010). The social sense: Susceptibility to others' beliefs in human infants and adults. *Science*, 330, 1830–1834. <https://doi.org/10.1126/science.1190792>.
- LaFreniere, P. J., & Dumas, J. E. (1996). Social competence and behavior evaluation in children ages 3 to 6 years: The short form (SCBE-30). *Psychological Assessment*, 8, 369–377. <https://doi.org/10.1037/1040-3590.8.4.369>.
- Leigh, J., Dettman, S., Dowell, R., & Briggs, R. (2013). Communication development in children who receive a cochlear implant by 12 months of age. *Otology & Neurotology*, 34, 443–450. <https://doi.org/10.1097/MAO.0b013e3182814d2c>.
- Lonigro, A., Laghi, F., Baiocco, R., & Baumgartner, E. (2013). Mind reading skills and empathy: Evidence for nice and nasty tom behaviours in school-aged children. *Journal of Child and Family Studies*, 23, 581–590. <https://doi.org/10.1007/s10826-013-9722-5>.
- 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, 622–646. <https://doi.org/10.1111/j.1467-8624.2007.01018.x>.
- 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 Developmental Disorders. *Journal of Autism and Developmental Disorders*, 29, 67–80. <https://doi.org/10.1023/A:1025922717020>.
- Nader-Grosbois, N., Houssa, M., & Mazzone, S. (2013). How could theory of mind contribute to the differentiation of social adjustment profiles of children with externalizing behavior disorders and children with intellectual disabilities? *Research in Developmental Disabilities*, 34, 2642–2660. <https://doi.org/10.1016/j.ridd.2013.05.010>.
- Pavarini, G., de Hollanda Souza, D., & Hawk, C. K. (2012). Parental practices and theory of mind development. *Journal of Child and Family Studies*, 22, 844–853. <https://doi.org/10.1007/s10826-012-9643-8>.
- Peskin, J., & Ardino, V. (2003). Representing the mental world in children's social behavior: Playing hide-and-seek and keeping a secret. *Social Development*, 12, 496–512. <https://doi.org/10.1111/1467-9507.00245>.
- Peterson, C. C. (2003). The social face of theory of mind: The development of concepts of emotion, desire, visual perspective, and false belief in deaf and hearing children. In B. Repacholi & V. Slaughter (Eds.), *Individual differences in theory of mind: Implications for typical and atypical development* (pp. 172–196). Psychology Press.
- Peterson, C. C., Garnett, M., Kelly, A., & Attwood, T. (2009). Everyday social and conversation applications of theory-of-mind understanding by children with autism-spectrum disorders or typical development. *European Child & Adolescent Psychiatry*, 18, 105–115. <https://doi.org/10.1007/s00787-008-0711-y>.
- Peterson, C. C., & Siegal, M. (2000). Insights into theory of mind from deafness and autism. *Mind & Language*, 15, 123–145. <https://doi.org/10.1111/1468-0017.00126>.
- Peterson, C. C., Slaughter, V. P., & Paynter, J. (2007). Social maturity and theory of mind in typically developing children and those on the autism spectrum. *Journal of Child Psychology and Psychiatry*, 48, 1243–1250. <https://doi.org/10.1111/j.1469-7610.2007.01810.x>.
- Peterson, C. C., & Wellman, H. M. (2009). From fancy to reason: Scaling deaf and hearing children's understanding of theory of mind and pretence. *British Journal of Developmental Psychology*, 27, 297–310. <https://doi.org/10.1348/026151008X299728>.
- Peterson, C. C., Wellman, H. M., & Liu, D. (2005). Steps in theory of mind development among children with autism, deafness or typical development. *Child Development*, 76, 502–517. <https://doi.org/10.1111/j.1467-8624.2005.00859.x>.
- Peterson, C. C., Wellman, H. M., & Slaughter, V. (2012). The mind behind the message: Advancing theory of mind scales for typically developing children, and those with deafness, autism, or Asperger syndrome. *Child Development*, 83, 469–485. <https://doi.org/10.1111/j.1467-8624.2011.01728.x>.
- Raven, J. C. (1938). *Progressive Matrices: A perceptual test of intelligence, individual form*. K.K. Lewis.
- Raver, C. C., & Leadbeater, B. J. (1993). The problem of the other in research on theory of mind and social development. *Human Development*, 36, 350–362. <https://doi.org/10.1159/000278223>.
- Roth, D., & Leslie, A. M. (1998). Solving belief problems: Toward a task analysis. *Cognition*, 66, 1–31. [https://doi.org/10.1016/S0010-0277\(98\)00005-5](https://doi.org/10.1016/S0010-0277(98)00005-5).
- San José Cáceres, A., Keren, N., Booth, R., & Happé, F. (2014). Assessing theory of mind nonverbally in those with intellectual disability and ASD: The penny hiding game. *Autism Research*, 7, 608–616. <https://doi.org/10.1002/aur.1405>.
- Senju, A., Southgate, V., Miura, Y., Matsui, T., Hasegawa, T., Tojo, Y., Osanai, H., & Csibra, G. (2010). Absence of spontaneous action anticipation by false belief attribution in children with autism spectrum disorder. *Development and Psychopathology*, 22, 353–360. <https://doi.org/10.1017/S0954579410000106>.
- Smogorzewska, J., Szumski, G., & Grygiel, P. (2019). The Children's Social Understanding Scale: An advanced analysis of a parent-report measure for assessing theory of mind in Polish children with and without disabilities. *Developmental Psychology*, 55, 835–845. <https://doi.org/10.1037/dev0000673>.
- Spek, A. A., Scholte, E. M., & Van Berckelaer-Onnes, I. A. (2010). Theory of mind in adults with HFA and Asperger syndrome. *Journal of Autism and Developmental Disorders*, 40, 280–289. <https://doi.org/10.1007/s10803-009-0860-y>.
- Tager-Flusberg, H. (1996). Brief report: Current theory and research on language and communication in autism. *Journal of Autism and Developmental Disorders*, 26, 169–172. <https://doi.org/10.1007/BF02172006>.
- Tager-Flusberg, H. (2007). Evaluating the theory-of-mind hypothesis of autism. *Current Directions in Psychological Science*, 16, 311–315. <https://doi.org/10.1111/j.1467-8721.2007.00527.x>.
- Tahiroglu, D., Moses, L. J., Carlson, S. M., Mahy, C. E., Olofson, E. L., & Sabbagh, M. A. (2014). The Children's Social Understanding Scale: Construction and validation of a parent-report measure for assessing individual differences in children's theories of mind. *Developmental Psychology*, 50, 2485–2497. <https://doi.org/10.1037/a0037914>.
- Tahiroglu, D., & Yagmurlu, B. (2014). *Turkish adaptation and psychometric properties of the first parent-report theory of mind scale*. [Poster presentation]. 18th National Congress of Psychology, Bursa, Turkey.
- Van Herwegen, J., Dimitriou, D., & Rundblad, G. (2013). Performance on verbal and low-verbal false belief tasks: Evidence from children with Williams syndrome. *Journal of Communication Disorders*, 46, 440–448. <https://doi.org/10.1016/j.jcomdis.2013.10.002>.
- Wellman, H. M. (2012). Theory of mind: Better methods, clearer findings, more development. *European Journal of Developmental*

- Psychology*, 9, 313–330. <https://doi.org/10.1080/17405629.2012.680297>.
- Wellman, H. M. (2014). *Making minds: How theory of mind develops*. Oxford University Press.
- Wellman, H. M., Cross, D., & Watson, J. (2001). Meta-analysis of theory-of-mind development: The truth about false belief. *Child Development*, 72, 655–684. <https://doi.org/10.1111/1467-8624.00304>.
- Wellman, H. M., & Liu, D. (2004). Scaling of theory-of-mind tasks. *Child Development*, 75, 523–541. <https://doi.org/10.1111/j.1467-8624.2004.00691.x>.
- Wellman, H. M., & Peterson, C. C. (2013). Theory of mind, development, and deafness. In S. Baron-Cohen, H. Tager-Flusberg, & M. V. Lombardo (Eds.), *Understanding other minds. Perspectives from developmental social neuroscience* (pp. 51–71). Oxford University Press.
- Wimmer, H., & Perner, J. (1983). Beliefs about beliefs: Representation and constraining function of wrong beliefs in young children's understanding of deception. *Cognition*, 13, 103–128. [https://doi.org/10.1016/0010-0277\(83\)90004-5](https://doi.org/10.1016/0010-0277(83)90004-5).
- Woolfe, T., Want, S. C., & Siegal, M. (2002). Signposts to development: Theory of mind in deaf children. *Child Development*, 73, 768–778. <https://doi.org/10.1111/1467-8624.00437>.
- Yagmurlu, B., Berument, S. K., & Celimli, S. (2005). The role of institution and home contexts in theory of mind development. *Journal of Applied Developmental Psychology*, 26, 521–537. <https://doi.org/10.1016/j.appdev.2005.06.00>.
- 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, 283–307. <https://doi.org/10.1037/0033-2909.124.3.283>.