



# The Development and Preliminary Psychometric Evaluation of the Theory of Mind Inventory: Self Report—Adult (ToMI:SR-Adult)

Tiffany L. Hutchins<sup>1</sup> · Laura Lewis<sup>2</sup> · Patricia A. Prelock<sup>1</sup> · Ashley Brien<sup>3</sup>

© Springer Science+Business Media, LLC, part of Springer Nature 2020

## Abstract

This study describes the development and psychometric evaluation of a new self-report measure of social cognition: the Theory of Mind Inventory:Self-report-Adult (ToMI:SR-Adult). Adults with autism (or a suspicion of autism;  $n = 111$ ) and typically developing adults ( $n = 109$ ) completed a demographic questionnaire and the ToMI:SR-Adult online. Both quantitative and qualitative self-reports of one's own theory of mind functioning were collected. The ToMI:SR-Adult performed well under all examinations of reliability and validity (internal consistency, accuracy of classification, contrasting-groups). The qualitative data confirmed impressions of validity and revealed that the adults in our sample had high levels of self-insight regarding their own theory of mind. The ToMI:SR-Adult is offered as a promising research and clinical tool for the assessment of social cognition in adults.

**Keywords** Autism · Assessment · Adult · Theory of mind · Social cognition · Self-report

Research in theory of mind (ToM) has been extremely active over the last 50 years and a wide variety of ToM tests have been developed. The earliest ToM assessments were direct measures of social cognition in the form of tests with ostensibly correct and incorrect answers. Although these measures found much use in the laboratory and clinic, one notable drawback was the observation of ceiling effects, especially when mentalizing was relatively good (e.g., Slaughter and Repacholi 2003). This led to the development of several 'advanced' ToM tests (e.g., Baron-Cohen et al. 2001; Devine and Hughes 2013; Happé 1994; Kaland et al. 2002; Muris et al. 1999; White et al. 2009) which represents innovation in assessment of social cognition although they also are accompanied by some serious limitations. Chief among these, the explicit nature of laboratory-type ToM tasks is quite unlike the ways that real life social dilemmas are presented (e.g., Hutchins et al. 2012, 2016). In fact, the

notion that task performance can exceed social cognitive functioning when it is applied in everyday life has been the topic of considerable concern (Astington 2003; Davies and Stone 2003; Klin et al. 2000). This raises important issues about the social validity of ToM assessment and recently, there has been renewed interest in the utility of self-report measures to complement existing tools measuring psychosocial characteristics in adults representing a wide range of clinical conditions.

This study reports on the development and preliminary psychometric evaluation of a new self-report measure to assess social cognition in adults with, or at risk for, ASD. A unique aspect of the Theory of Mind Inventory: Self Report—Adult (ToMI:SR-Adult) is that items were designed to characterize how autistic individuals experience perspective-taking opportunities, rather than relying on test performance or the impressions of other informants. In validation studies of an earlier, parent-informant version of the Theory of Mind Inventory (ToMI; Hutchins et al. 2008, 2012), typically developing samples evidenced ceiling effects in late childhood but autistic individuals who were older and those with relatively good mentalizing skills did not approach the ceiling. Subsequently, a similar pattern (i.e., no ceiling effects) was observed for highly verbal autistic adolescents (Lerner et al. 2011) suggesting that this tool might be appropriate for detecting even the most subtle

✉ Tiffany L. Hutchins  
Tiffany.Hutchins@med.uvm.edu

<sup>1</sup> Department of Communication Sciences and Disorders,  
University of Vermont, 407 Pomeroy Hall, 489 Main Street,  
Burlington, VT 05405, USA

<sup>2</sup> Department of Nursing, University of Vermont, Burlington,  
USA

<sup>3</sup> University of Vermont, Burlington, USA

theory of mind challenges in older samples. More recently, the criterion-related validity of the Theory of Mind Inventory-2nd Edition (ToMI-2) was evaluated as a self-report tool for autistic adults (Crehan et al. 2019). In that study, ToMI-2 Advanced Subscale scores were positively correlated with self-reported social responsiveness and negatively correlated with self-reported depressive symptoms and more typical visual scan patterns in response to social stimuli: all of these are theoretically-, empirically-, and clinically-relevant links that provide support for the development and evaluation of a self-report for autistic adults as an indicator of advanced social cognition.

Although self-report is currently being used to assess a variety of characteristics in ASD (e.g., anxiety, depression, systemizing, internalizing), self-assessment of social cognition remains a woefully understudied topic.

Still, research in this area is critical as more adults are seeking diagnostic evaluations, resulting in a need for tools that are accessible in situations when parent-report cannot (or will not) be available. Moreover, targeted and sensitive measurement of one's own theory of mind strengths and challenges would enable service providers to not only adapt social skills supports and monitor treatment progress in an ecologically valid way, but may also provide the 'buy in' from clients crucial to intervention success" (Crehan et al. 2019, p. 659).

Of course, assessing one's own ability to perspective-take is a potentially confounding concept. "If you have less developed theory of mind, would you not rate yourself higher on a theory of mind measure, as you would have limited insight in this area?" (Crehan et al. 2019, p. 663). Indeed, persons with ASD evidence higher rates of alexithymia (i.e., an inability to describe one's own feelings), may experience difficulty with self-introspection, and are limited in their ability to read the minds of others (i.e., which are all constructs that the ToMI:SR-Adult is intended to assess). Nonetheless, self-insight into psychosocial functioning is also well documented in ASD (Dritschel et al. 2010; Schriber et al. 2014). In fact, autism researchers and related professionals are now challenging the traditional notion that autistic individuals lack awareness of their own social-cognitive difficulties. "By recognizing self-report as a valid viewpoint, researchers and clinicians are gaining insight into how people experience autism, rather than relying only on others' accounts and observations" (Yuhas 2018, p. 2). Said another way, for some autistic persons, ASD may contribute to difficulties in self-insight, but autism *does not preclude* insight, awareness, and the ability to self-report. In fact, we will argue that many autistic adults possess high degrees of social awareness and are uniquely situated to report on their own ToM challenges. This is important because self-report helps move us toward

assessment of ToM that is socially valid, client-centered, and meaningful in everyday life. Moreover, capitalizing on the insights gleaned through self-report can help us begin a conversation with clients about their perceived strengths and challenges, their specific priorities for intervention, and their preferred treatment strategies and supports.

## Purpose

Our primary goal was to develop a content valid index of ToM that would be sensitive to variation in a wide range of *advanced* ToM understandings: those that can represent the subtlest and difficult-to-detect domains, especially among those with high cognitive and language abilities. An embedded instrument development and validation variant of the embedded mixed methods design was used to develop and pilot the ToMI-SR-Adult, with the quantitative strand taking priority and the qualitative strand supplementing refinement and evaluation of the tool (Creswell and Plano-Clark 2017). Quantitative and qualitative strands occurred concurrently.

## Method

### Item Development and Content Validity

Content validity refers to whether the items on a test sufficiently tap the various aspects of a construct (content coverage) while avoiding content unrelated to the construct (construct relevance). In the test item development phase, this typically involves a panel of experts (McCauley 2001). The content of the ToMI:SR-Adult was developed by three experts with extensive experience researching ASDs. Two of the three experts also developed several previously published ToM assessments and have over 15 years of expertise in defining and applying ToM concepts to ASD populations.

We defined advanced ToM domains as those that were later developing in typically developing populations (specifically, the school aged and beyond). This was driven by the aforementioned studies establishing a lack of ceiling effects for these domains in autistic adolescents and adults (Crehan et al. 2019; Lerner et al. 2011). Of course, all forms of validity (including content validity) are relative and contextual judgments based on an accumulation of evidence. Moreover, content validity must be evaluated in light of the *goals of assessment*. In this case, recall that our primary goal was to develop a measure that would assess a wide range of ToM domains that could inform research and clinical practice. In addition, it is important to acknowledge that the term ToM has been in transition over the last few decades. Although once used in a narrow sense (to refer only to the ability to pass tasks of false

belief), the term ToM “refuses to be corralled” (Astington and Baird 2005, p. 4) and is considered by many contemporary researchers as a synonym for social cognition (Hutchins et al. 2012) which refers to all social reasoning processes in one’s self and others. As such, the content domain of ToM is bigger than immense.

Characterizing the scope of ToM is further complicated by the nature of ToM itself: ToM is *developmentally and functionally intertwined* and massively interactive with a range of other human-specific cognitive achievements. These include autobiographical (specifically episodic) memory, future thinking, time perception, emotional introspection, self-concept, and the development of a coherent identity to name a few (e.g., Allman and Faltar 2015, Brien et al. in press; Nelson 2003; Spreng and Andrews-Hanna 2015; Spreng et al. 2008). As such, we consider these domains as rightfully belonging to the construct of ToM as they affect and are affected by cognitive compensations in the default network of which ToM is a part. They are also disrupted in ASD and are important for adaptive functioning. For these reasons, such domains were considered for item inclusion.

The original pilot version of the ToMI:SR-Adult was expansive. It consisted of 122 items designed to tap a wide range of advanced social cognitive understandings, the content of which was guided by the immense theoretical and empirical research base in ToM in both typically developing adolescents and adults and those with ASD. In addition, we often developed multiple items for a single construct (e.g., episodic memory, empathy) with slightly different wording so we could determine the best way to ask the question. This resulted in redundancy in the original pilot version of the measure and so our instructions included language that explained the reason for the repetitive content. In addition, the pilot version of the ToMI:SR-Adult offered respondents the opportunity to comment on any of the items, specifically to “share your ideas about a question or your reasoning behind your response.”

From the initial pool of 122 items, 60 were retained in the final version of the measure. These items were chosen on the basis of their psychometric properties and their ability to tap the broadest content relevant to the construct of advanced ToM. Over a 6-month period during which data were collected, the authors met to review the statistical properties of each item as well as any comments from respondents so as to identify good and poor performing items. The 60 items retained were those that (1) demonstrated the greatest mean difference between ASD and non-ASD groups (as reported below, all were significantly different at  $p < 0.001$  with large effect sizes), (2) appeared to be clear, meaningful, and face valid indicators on the basis of Cronbach’s alpha and respondents’ qualitative comments and, (3) best contributed to the content coverage of the expansive construct that is advanced ToM as judged by the expert developers.

Following the identification of the final 60 items, exploratory principle components analysis was conducted to examine the dimensionality of the measure. No factors (subscales) were identified using either Varimax or oblique rotation suggesting that the items tapped a unitary construct (presumably advanced theory of mind).

The 60 items comprising the ToMI:SR-Adult and the constructs they are intended to assess are presented in the Table 1. Each item takes the form of a statement (e.g., “If I had a friend who was sad about something, I would feel sad myself”) and is accompanied by a 20-unit continuum anchored by ‘definitely not’, ‘probably not’, ‘undecided’, ‘probably’, and ‘definitely.’ The respondent is asked to read a statement and place a hash mark at the appropriate point along the continuum. The continuum and hash mark response arrangement was favored over a more traditional Likert-type scale for its ability to be sensitive to values between anchors and therefore enhance precision: a proposition that was empirically demonstrated on the earlier parent-informant version of the measure (Hutchins et al. 2008, 2012). Still, the response arrangement was not expected to be familiar to the adults who would be self-reporting and confusion about how to respond could occur. For this reason, the instructions included examples to make clear what is considered correct and incorrect forms of responding.

A formal analysis was conducted to determine the reading level of the ToMI:SR-Adult. According to the Flesch-Kincaid readability index (Kincaid et al. 1988), the ToMI:SR-Adult received a score of 82.5 (scores range from 0 to 100 with higher scores indicating easier reading) which is equivalent to a reading level of grade 5 which should be easily understood by the adult population for which this measure is intended. We adopted, however, a more conservative estimate and formally recommend that respondents have a reading level of no less than the 8th grade.

## Qualitative Analysis

Qualitative data were collected with the aim of understanding how participants interpreted the wording and content of individual items. A secondary purpose of this qualitative strand was to explore the validity of the self-report aspect of the tool by gaining information about self-awareness, personal insight, and perspective-taking as they relate to theory of mind.

Braun and Clarke’s (2006) method for thematic analysis was used to analyze qualitative data collected from autistic participants. In this flexible qualitative method, researchers identify patterns and themes from within the data using a six-phase approach, including: familiarizing yourself with your data; generating initial codes; searching for themes; reviewing themes; defining and naming themes; and producing the report. The first two

**Table 1** Item content and the intended construct of each item comprising the ToMI:SR-Adult

Item	Construct
1. If I had a friend who was sad about something, I would feel sad myself	Affective empathy
2. I understand that it is possible to experience two conflicting emotions at the same time (e.g., being sad that a sick pet died but being happy that it is no longer in pain)	Mixed emotions
3. I can accurately judge the passage of time (e.g., I am good at knowing when 5 min, 30 min, or 2 h has gone by)	Time perception
4. I have rich memories for my own past experiences (when recalling something, I remember where I was, when something happened, who else was there, what I was thinking or feeling)	Episodic memory
5. I understand the metaphor in the following sentence: "After many years of working at the jail, the prison guard had become a hard rock that could not be moved"	Nonliteral language: metaphor
6. If I heard a waitress say to a coworker "Hey, the ham sandwich left me a big tip!", I would understand that the waitress was referring to a customer who had ordered a ham sandwich	Nonliteral language: metonymy
7. If I drove up to my friend's house in a big car and my friend said "I didn't know you were a bus driver", I would understand that my friend was commenting on the size of the car	Nonliteral language: metaphor
8. When I look at this image, I immediately see what looks like a person	Social perception
9. A 'double bluff' is a type of deception where someone tries to deceive another person by actually telling the truth. I understand what double bluff is and why people use it	Double bluff
10. I understand why people often cling to mistaken beliefs despite overwhelming evidence to the contrary	Motivated reasoning
11. I understand that people can hold two or more contradictory beliefs, ideas or values at the same time (e.g., believing in fuel efficient cars but wanting a large fuel inefficient car)	Hypocrisy
12. I understand that people in a community share common notions that are considered obvious and known to all (e.g., people in our culture tend to assume that all children like sugary treats)	Social common sense
13. I understand how people make snap judgments about others based on very little information (e.g., Justin knows Bradley is wealthy so Justin assumes Bradley is 'out of touch' and unable to understand regular people)	Stereotypical thinking
14. I understand that when it comes to getting along with others, honesty is not always the best policy	White lie appreciation
15. I understand how people use facial expressions to hide how they really feel (e.g., people can smile when they are sad and laugh when they feel embarrassed)	Display rules
16. I can tell the difference between when a friend is teasing in a nice way and when a bully is making fun of someone in a mean way	Complex social judgment
17. I can tell the difference between lies and jokes	Lies v. jokes
18. Sometimes people do things deliberately and sometimes they make mistakes or do things by accident. I can tell when people do things on purpose or by accident	Intentionality
19. People do certain things when they are not interested in talking to us (e.g., they might look away for a long time or start fidgeting). I can recognize when a listener is not interested in what I am saying	Complex social judgment
20. People do certain things when they are disgusted (e.g., they might squint their eyes, crinkle their nose, or slowly shake their head). I can recognize when others are disgusted	Emotion recognition: disgust
21. I can put myself in other people's shoes and understand how they feel	Empathy (cognitive and affective)
22. When I hear jokes like "What is black, white, and <i>read</i> all over? It's a newspaper!" or "The duck said to the bartender, 'put it on my <i>bill</i> '", I immediately understand the humor in this play on words	Humor: play on words
23. I am good at getting the jokes that other people tell	Humor appreciation
24. I speak differently to young children compared to adults (e.g., I use simpler language when speaking to youngsters)	Audience adaptation
25. People do certain things when they are sad (e.g., they might look downward, have a frown, or talk in a soft voice). I can recognize when others are sad	Emotion recognition: sad
26. People do certain things when they are scared (e.g., they might raise their eyebrows, lean backwards, or talk in a high-pitched voice). I can recognize when others are scared	Emotion recognition: scared
27. People do certain things when they are surprised (e.g., they might open their eyes very wide or open their mouths really wide). I can recognize when others are surprised	Emotion recognition: surprise
28. I understand the kinds of situations that would make someone feel embarrassed	Situation-based emotion: embarrassed
29. I can accurately identify and reflect on my own emotions (that is, I know what I am feeling and why I am feeling it)	Introspection: emotion

**Table 1** (continued)

Item	Construct
30. I can predict my own emotions to better plan for the future (e.g., if going on a long car ride, I know that I might get bored so I might bring something to do in the car)	Future thinking
31. I understand the kinds of situations that would make someone feel proud	Situation-based emotion: pride
32. I can predict people's emotions based on the situation they are in	Situation-based emotion
33. I am good at planning for the future	Future thinking
34. I can predict how my words and actions will be interpreted by others	Social awareness
35. I can read other people's body language to tell how they feel (e.g., feeling happy, sad, mad, or scared)	Social perception: body language
36. I have good eye contact when I am talking to people	Eye contact
37. I understand how people can move their bodies to mislead others (e.g., someone who 'walks tall' to hide a lack of confidence)	Display rules
38. I can read the tone of other people's voices to tell how they feel (e.g., I can tell when others are excited, angry, or bored)	Social perception: voices
39. I can read people to tell if they like me or not	Social perception: affinity
40. I understand the kinds of situations that would make someone feel jealous	Situation-based emotion: jealousy
41. I can tell when others are being insincere	Lie detection
42. I can not only remember but also re-experience or relive an earlier experience (e.g., when thinking of a birthday party, I can re-experience the sights, smells, or sounds)	Episodic Memory
43. I feel sad or bad when I think about others who experience misfortune	Sympathy
44. I understand the kinds of situations that would make a person feel two different emotions at the same time (e.g., being mad and sad at the same time or being happy and excited at the same time)	Emotion blends
45. I understand verbal irony. For example, if it were raining and someone said "Looks like a really nice day outside", I would understand that they didn't actually think it was a nice day	Nonliteral language: verbal irony
46. I understand my own actions (that is, I know what I am doing and why I am doing it)	Introspection: action
47. I understand my own desires (that is, I know what I want and why I want it)	Introspection: desire
48. I understand my own plans and goals for the future (that is, I know what I will do and why I will do it)	Introspection: planning
49. If I were watching a movie with a friend, I could predict what my friend was thinking and feeling in response to the movie	Cognitive empathy
50. I am good at explaining my feelings	Emotional intelligence
51. I can imagine myself in the future and see what I am doing, thinking, and feeling	Future thinking
52. When I look at people's faces, I tend to look at their eyes because the eyes give me good information about what someone is thinking or feeling	Social perception: reading eyes
53. I am good at picking up on other people's moods without having to think about it	Social perception: mood
54. I am good at predicting events in terms of whether they are likely or unlikely to actually occur	Prognostication
55. I understand the metaphors people use (e.g., "That man is a lion" or calling a puddle of oil a "dead rainbow")	Non-literal language: metaphor
56. I am socially aware (that is, I am aware of and understand the social features of the situation I am in)	Social awareness
57. I can shift a topic of conversation with skill	Expressive pragmatics: conversation
58. When I recount stories, I talk about people's thoughts and feelings to explain their actions	Expressive pragmatics: narrative construction
59. I understand proverbs even if they are unfamiliar to me (e.g., "Envy is destroyed by true friendship")	Nonliteral language: proverbs
60. I am good at beginning and ending conversations	Expressive pragmatics: conversation

authors conducted the first three phases independently and then discussed findings to reach consensus on themes most "key" to the research question. The second author

conducted the final three phases and the final report was reviewed by all three authors to ensure adequate representation of data.

We approached this thematic analysis from a realist theoretical lens. That is, we contend that the construct of theory of mind exists regardless of how it is interpreted or understood in individual experiences. In this study, our aim was not to capture how participants defined or made meaning of their experiences of theory of mind, but instead to capture how they related to concepts that are core to the pre-defined construct of theory of mind.

Themes were defined as patterns within the data that were “key” to understanding the research purpose, not necessarily based on quantifiable prevalence within the dataset. Notably, though, themes identified were present across nearly all items and represent a vast majority of responses. We identified themes at the latent level, prioritizing the underlying ideas and conceptualizations over the semantic content of data. For example, in a comment related to an item capturing time perception (item 3), one participant commented, “I don’t notice time when I’m hyper focused, and I’m hyper focused more often than not.” This item was interpreted to shed light on the fact that this individual could identify his or her own ability to detect passage of time, not to explore how hyper-focus related to his or her perceived passage of time.

### Standardization Sample

An international sample of 300 self-identified fluent English speakers (111 with known or suspected ASD and 189 without ASD) completed the ToMI:SR-Adult online via a secure survey platform. From the total sample ( $n = 300$ ), 80 cases were systematically removed to create ASD and non-ASD groups matched on sex assigned at birth and socioeconomic status (assessed via reports of gross annual combined household income). This is important as both female sex (particularly with regard to the advanced social cognitive abilities of typically developing samples; Baron-Cohen 2009; Bosacki 2000; Devine and Hughes 2013) and higher socioeconomic status (SES; Cutting and Dunn 1999; Hughes et al. 2005; Pears and Moses 2003) are positively associated with theory of mind development.

### ASD Sample

As noted previously, ASD and non-ASD groups were distribution matched on SES ( $p = 0.39$ ). For the ASD group, SES (“combined annual household income before taxes”) ranged from \$0.00–\$750,000.00 ( $M = \$66,449.00$ ;  $SD = \$74,210.00$ ). The ASD group included 111 adults (56 males ages 18–72,  $M = 31.68$ ,  $SD = 13.50$ ; 55 female ages 18–72,  $M = 34.84$ ,  $SD = 13.90$ ) from nine countries/regions (primarily United States, United Kingdom, Europe, and Australia). To be included in the ASD group, respondents had to report that they had a formal diagnosis *or suspicion of autism*. This was important because delayed diagnosis of ASD into adulthood is common, and self-diagnosis is a growing phenomenon. In fact, there is a growing body of evidence on practical, developmental, clinical, and autism-trait-related barriers that prevent diagnosis for adults on the spectrum, leading to a “lost generation” of adults with more subtle symptomology (Lai and Baron-Cohen 2015; Lewis 2017). We did not want barriers such as access to assessment or health care providers with sufficient knowledge about ASD in adults to impact participation of a self-recognized population. We also imagine that, ultimately, the ToMI:SR-Adult can be used as a tool to aid in the identification of highly verbal adults on the autism spectrum who are sometimes difficult to diagnosis through the use of existing measures. A breakdown of how respondents in the ASD group evaluated themselves with regard to diagnosis is presented in Table 2. A one-way ANOVA was conducted to explore whether the composite ToMI:SR-Adult score varied by the four methods of evaluation. Results revealed no significant differences (omnibus and post-hoc comparisons were equal to or exceeded  $p > 0.26$ ).

Although precise estimates are difficult to obtain, the prevalence of complex clinical comorbidity is well-documented such that individuals with ASD often meet criteria for at least one additional psychiatric disorder (e.g., Mosner et al. 2019; Rosen et al. 2018). For this reason, we expected to find high rates of psychiatric comorbidity in our ASD sample and did not attempt to match the control group on this basis. Rather, we describe the major classes of co-occurring conditions in our sample and explored whether such

**Table 2** ASD reported evaluation method for presence of ASD

How evaluated	ASD; N = 111	
	Frequency	%
I am self-diagnosed	15	13.5
I have received a formal diagnosis	63	56.8
General practitioner or mental health professional believes I have ASD but I have not been formally assessed	22	19.8
Suspect I may be autistic but would not characterize myself as ‘self-diagnosed’ at this time	11	9.9



conditions were associated with differences in the ToMI:SR-Adult scores. Conditions were reported representing six categories (ADHD, speech-language disorder, intellectual disability, developmental delay, hearing impairment, other problems) of known or suspected concomitant disorders (see Table 3). A series of one-way ANOVAs comparing those with and without each co-occurring condition revealed no significant differences (all comparisons  $p > 0.33$ ).

### Non-ASD sample

The non-ASD group (N-ASD) included 109 adults (54 males ages 21–85,  $M = 46.37$ ,  $SD = 18.55$ ; 55 females ages 18–78,  $M = 33.09$ ,  $SD = 14.14$ ) from five countries/regions. For the N-ASD group, SES (“combined annual household income before taxes”) ranged from \$0.00 to \$320,000.00 ( $M = \$96,548.50$ ;  $SD = \$37,222.00$ ).

## Results

### Quantitative Analyses

#### Reliability

Reliability refers to the characteristics of dependability of measurement and is a necessary, but not sufficient criterion for validity (McCauley 2001). Reliability of the ToMI:SR-Adult was examined by assessing internal consistency: a measure of homogeneity of content. When internal consistency is high, this is generally taken as evidence that the items on a measure tap a unitary construct (in this case presumably, ToM). Cronbach’s alpha was calculated for the combined (ASD and N-ASD) group resulting in a value of 0.98. Alpha was not improved by removing any items, suggesting that all were contributing to the assessment of a unitary construct. When the groups were examined separately, the ASD group evidenced an alpha of 0.97

and the N-ASD group evidenced an alpha of 0.95, again indicating a high degree of internal consistency.

#### Validity

Construct validity refers to the extent to which a measure actually measures what it intends to measure. We explored several kinds of validity including content validity, contrasting-groups validity, and accuracy of classification.

#### Content Validity

Typically, the first step in establishing validity is logical and theoretical as opposed to statistical (McCauley 2001). Content validity refers to the degree to which items on a measure adequately tap the construct of interest while avoiding irrelevant content (McCauley 2001). The content validity of the ToMI:SR-Adult was carefully considered during the item development phase and in the item-revision of the original 122-item version of the measure (described earlier). The developers of the ToMI:SR-Adult who collaborated on item development and revision were experts in ASD, theory of mind, and the assessment of social cognition in ASD.

#### Contrasting-Groups Validity

One examination of validity employed a contrasting-groups method of construct validation. Specifically, we reasoned that a construct valid measure of advanced ToM competence should distinguish ASD and N-ASD groups. A total of 61 independent samples t-tests were conducted to examine group differences at the composite- and item-level. Results for all comparisons resulted in significant between-group differences for all 61 comparisons at  $p < 0.001$ . Moreover, Cohen’s  $d$  effect size estimates ranged from 0.60 to 1.86 (mean = 1.11; median = 1.06) which are medium to very large.

In addition, we expected a construct valid measure of advanced social cognition to be sensitive to a female advantage in social cognition that has been documented in N-ASD samples in several advanced and late-developing theory of mind domains. In line with this expectation, a significant difference was found for the composite score for the N-ASD sample such that females ( $M = 17.75$ ;  $SD = 1.45$ ) obtained higher scores than did males ( $M = 16.41$ ;  $SD = 2.00$ ),  $t(107) = 3.99$ ,  $p < 0.01$ . No female advantage was observed for the ASD sample with females ( $M = 11.93$ ;  $SD = 3.35$ ) and males ( $M = 11.72$ ;  $SD = 3.42$ ) obtaining similar scores,  $p = 0.71$ .

**Table 3** ASD and N-ASD sample reported (known or suspected) clinical conditions

Concomitant conditions (known or suspected)	ASD sample; N = 111		N-ASD sample; N = 109	
	Frequency	%	Frequency	%
Attention deficit hyperactivity disorder (ADHD)	32	28.8	6	5.5
Speech language disorder	24	21.6	5	4.6
Intellectual disability	50	45.0	12	11.0
Developmental delay	25	22.5	0	0
Hearing impairment	13	11.7	12	11.0
Other problem	47	42.3	17	15.6

## Accuracy of Classification

The data sets described above were used to calculate sensitivity, specificity, positive predictive validity, negative predictive validity, and overall accuracy. This is important as an additional test of construct validity and for the potential of the ToMI:SR-Adult for use in clinical practice. Given the moderate female advantage noted in our N-ASD sample, separate analyses for males and females was conducted to identify the optimal cut score for each population. To achieve this, histograms for males with and without ASD (or suspected ASD) and histograms for females with and without ASD (or suspected ASD) were examined separately. The optimal cut-score (i.e., the score that yielded the highest number of correct decisions) for males was 15.0 (equivalent to the 21<sup>st</sup> percentile relative to N-ASD males) and for females was 16.5 (equivalent to the 18.5<sup>th</sup> percentile relative to N-ASD females). Crucially, all signal detection analyses were initially conducted separately for those who were formally identified with ASD (i.e., had a diagnosis) and those who reported suspicion of ASD, however, the results for these groups were almost identical. For this reason (and to ensure a larger, more reliable sample), we collapsed those groups for the following analyses.

**Sensitivity** Sensitivity refers to the correct detection rate or the probability of a test to give a positive result when the person being assessed truly does have the disorder (McCaulley 2001). Our data reveal that 92.72% of females with known or suspected ASD were captured by the ToMI:SR-Adult when composite scores fell below a cut score of 16.50 when compared to a N-ASD normative sample. Similarly, 82.14% of males with known or suspected ASD were captured by the ToMI:SR-Adult when composite scores fell below a cut score of 15.00 when compared to a N-ASD normative sample.

**Specificity** Specificity refers to correct rejection rate or the probability of a test to give a negative result when the person being assessed truly does not have the disorder (McCaulley 2001). Analyses for specificity revealed that 81.81% of N-ASD females were correctly rejected when composite scores were greater than or equal to a cut score of 16.50 and 79.63% of N-ASD males were correctly rejected using a cut-score of 15.00.

**Positive and Negative Predictive Validity** Positive predictive validity refers to the probability that the condition is present when the test is positive. Positive predictive validity was 83.61% for females and 80.70% for males. Negative predictive validity is the probability that the condition is not present when the test is negative. Negative predictive validity was 91.84% for females and 81.13% for males.

**Table 4** Rows by column contingency table for accuracy of classification for females

ToMI:SR-Adult cut score = 16.5	Diagnosis (or suspected diagnosis) of ASD	
	Yes	No
Yes	51	10
No	4	45

Sensitivity = 92.72%  
Specificity = 81.81%  
Positive predictive validity = 83.61%  
Negative predictive validity = 91.84%  
Overall accuracy: = 87.27%

**Table 5** Rows by column contingency table for accuracy of classification for males

ToMI:SR-Adult cut score = 15.0	Diagnosis (or suspected diagnosis) of ASD	
	Yes	No
Yes	46	11
No	10	43

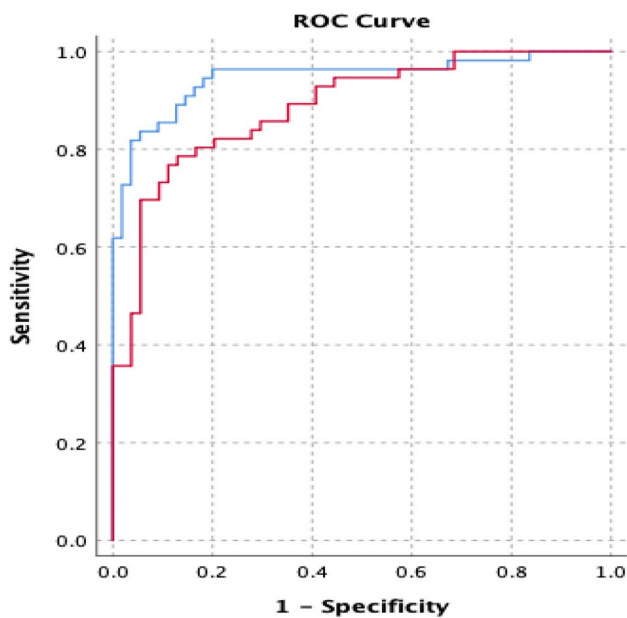
Sensitivity = 82.14%  
Specificity = 79.63%  
Positive predictive validity = 80.70%  
Negative predictive validity = 81.13%  
Overall accuracy: = 80.90%

**Overall Accuracy** The overall accuracy of the ToMI:SR-Adult for identification of ASD using a cut score of 16.50 for females (equivalent to approximately the 18th percentile for N-ASD females) and 15.00 for males (equivalent to the 21st percentile for N-ASD males) was 87.27% and 80.90%, respectively. The Rows by Column contingency table for sensitivity, specificity, positive predictive validity, negative predictive validity, and overall accuracy are represented in Tables 4 and 5 for females and males respectively.

## Receiver Operating Characteristics

We also examined receiver operating characteristics and area under the curve (ROC/AUC) to explore the effectiveness of the ToMI:SR-Adult for detecting clinical levels of social cognitive dysfunction (as would be observed in those with or suspected of having ASD). The ROC/AUC curve “is a





**Fig. 1** ROC curve for females (blue) and males (red)

measure of the overall performance of a diagnostic test and is interpreted as the average value of sensitivity of all possible values of specificity” (Park et al. 2004, p. 13). The ROC/AUC curve and associated statistics are presented in Fig. 1. The ToMI:SR-Adult achieved a value of 0.95 for females and 0.89 for males and an overall value of 0.92 which is considered excellent (Compton et al. 2006).

## Qualitative Results

Four qualitative themes were identified as key to understanding the validity of a self-report measure of theory of mind. First, participants provided rich descriptions of the quality of substantive problems they experienced related to theory of mind. Second, they were able to improve specific theory of mind skills with self-awareness and practice. Third, they frequently described understanding concepts on a theoretical level but struggled with spontaneous and situationally appropriate execution of skills. Fourth, there was wide variability in domains that were perceived as strengths versus weaknesses among this sample. Overall, participants’ responses indicated they were highly aware of, and had much insight into, their own theory of mind challenges. Themes that emerged are explored in depth in the following sections.

### Rich Descriptions of Quality of Experiences

The comments of respondents provided thoughtful reflections on how individuals experienced specific theory of mind tasks and skills. For example, in response to an item on episodic memory (item 4), one participant shared, “My

memories seem more organized around ideas than around events! I might have snippets of such ‘richness’ with experiences but that is rare... However, I’m good at understanding complex systems linked by a mix of logic and emotion and (sensory stuff?).” Another participant commented on his or her ability to accurately identify and reflect on emotions (introspection: emotion, item 29) and shared:

Sometimes I know how it feels, but I don’t know what it is. I could describe the feeling, the sensorial part, but not address a name or the origin. Or I can tell why it happened, after what occurred, or what someone said, but I don’t give me permission or don’t know if it’s okay or not. It’s like lacking permission or approval of someone else to feel the way I feel. (\*By the way, this questionnaire is being a therapy so far. Thank you, dear researchers).”

Importantly, participants showed evidence of an ability to extrapolate meaning from their experiences and apply those when responding quantitatively to items on the instrument. For example, when asked about the ability to plan for the future (future thinking, item 33), one participant shared, “I am sixty-six and went bankrupt at 62... nuff said.” Another participant provided an example of a time they misunderstood verbal irony in a natural setting to illustrate a weakness in this skill (nonliteral language: verbal irony, item 45). In response to an item on mixed emotions (item 2), one participant shared, “Boy, did this hit home... at my Mom’s funeral I was happy she wasn’t suffering any more. My family took it as I was happy she died. I wasn’t. I was happy for her suffering ending.” In all of these cases, participants pulled examples from their own lived experiences and applied them in their interpretation of their skill level. Participants consistently demonstrated a detailed awareness of how their experiences matched or differed from items on the instrument.

### Progression of Skills Over Time

Comments suggested that participants were able to improve many theory of mind skills over time, often describing items on the instrument as “learned skills.” Participants remarked on their personal growth, sharing comments such as, “I would say I am MORE socially aware than I was 10 years ago!” (social awareness, item 56), and, “This is something that developed in me more recently,” (sympathy, item 43), and, “I started doing this during my 40’s” (social perception: reading eyes, item 52).

They often cited practicing or receiving training in a particular skill, such as through therapy. In response to an item on expressive pragmatics in conversation (item 57), for example, participants commented, “A skill I have learned, but a skill nonetheless,” and, “Years of education and training.” Others shared, “I have had forty plus years of therapy

so I have been trained in this subject” (introspection: emotion, item 29), and, “I am good at this because of years of practice, lots of therapy, and extensive reading on the subject” (emotional intelligence, item 50). Another commented, “Readings on socialization, sociology, and structuralism helped me in making this realization” (social common sense, item 12).

Even with practice, however, participants shared that these skills often still felt unnatural or uncomfortable, or that their behavioral response did not match their underlying mentality. For example, in response to an item on recognizing when a listener is not interested, one participant shared, “I CAN, but I usually don’t do what I expect myself to do, which is to shut-off and leave. I look, I see, those body sensations come to me like a radar beeping ‘something’s not right’, but I don’t take any action” (complex social judgment, item 19). Another shared about eye contact (item 36), “I look right past their ear so that while others think I am making eye contact I’m really not.” Participants often described learning either the underlying mentality or the applied behavior associated with a particular skill, but without both pieces, they still perceived a weakness in the skill overall.

### Challenges with Real-World Applicability

Across items, participants indicated that they often had theoretical understanding of a skill but struggled to spontaneously apply this skill in real-world scenarios. For an item on white lie appreciation (item 14), for example, one participant responded, “Theoretically I get it but I don’t know how to apply it in real life.” Another commented on an item on metonymy (item 6), “I understand it reading it, but I’m not sure I would get it if I heard it out loud.” Participants described processing skills at a conscious level, rather than executing these skills as a natural response. For instance, one participant described his social awareness (item 56): “I would say I do it in software rather than in hardware, there. That is, I have to analyze the situation with reason and experience; it doesn’t come to me unconsciously like it does for some.” Similarly, in an item related to double bluff (item 9), one participant shared, “I have heard of the truth being more unbelievable than a lie so I understand the principle, but actually applying it or why someone would use it as a ‘bluff’ is beyond me.”

Participants often shared that they could execute a particular skill with time, but not immediately. For example, on an item about ability to pick up on other people’s moods without having to think about it (social perception: mood, item 53), several participants shared, “I have to think about it.” On an item about differentiating between friendly teasing and when a bully is making fun of someone in a mean way (complex social judgment, item 16), one participant shared, “Often I cannot do this well in real time. If I think about

the interaction afterward, I can figure it out.” Other skills required significant effort, such as eye contact (item 36): “I can put the effort in to do this,” and, “I can force myself to make eye contact, but it takes a lot of energy and I can’t do it for too long.”

Participants also frequently described understanding that another person might respond a certain way in a given situation, but not understanding *why* or *how* that person responded in that way. For example, in an item on using facial expressions to hide real feelings (display rules, item 15), one participant commented, “I understand why, but I don’t understand how they manage to mask them. I have a hard time pretending to feel a way I certainly do not feel, even if I remain guarded of my emotions.” In relation to jealousy (item 40), one participant wrote, “Other people get jealous more easily than I do. I understand certain kinds of jealousy (like when your best friend suddenly wants to spend all their time with a new friend), but not other kinds (like meeting someone who’s really nice and smart and has their life together).” In providing this example, the participant indicates a theoretical understanding of situations that may cause others to become jealous, but does not personally understand the full sweep of this experience.

### Variability in Strengths Versus Weaknesses

There was wide variability in participant reports of strengths versus weaknesses in specific theory of mind skills, both across items and within items. Participants frequently commented that one item was a strength for them while other participants reported that same item as an area of weakness. For example, when asked about future thinking (item 30), one participant shared, “I struggle with this,” while another shared, “This has become somewhat of a strength for me.” On an item related to understanding unfamiliar proverbs (nonliteral language: proverbs, item 59), one participant responded, “A particular strength of mine” while another shared, “Only if they are familiar or explained to me.” On an item about sympathy (item 43), some participants indicated they experienced extreme levels of sympathy, for example, “Too much at times,” while others indicated a lack of sympathy, for example, “I get numb to this so it doesn’t always register.” Similar variability was found across items in our measure.

Comments also shed light on variability of individual experiences, even when participants agreed that an item was a problem area for them. For example, on an item related to the ability to begin and end conversations (expressive pragmatics: conversation, item 60), most participants shared comments such as, “Can’t start but if I do I won’t quit...,” and, “Good at beginning, but not ending.” Yet, another participant shared, “Much better at ending them than starting them.” In an item on eye contact (item 36), some participants

shared a challenge with using too much eye contact, for example, “Sometimes I don’t know when to break eye contact,” and, “I tend to hold eye contact way too long, my wife says it makes me come across as intense.” Others indicated that they avoided eye contact altogether, saying, “Try to avoid eyes, but will look to the whole face or to the eyes area.” While all of these participants rated this domain as a weakness, the specific challenges varied across participants.

## Discussion

Most existing measures of theory of mind lack an examination of how individuals experience theory of mind dilemmas in real-world social situations thus raising serious questions related to social and ecological validity. The data reported here for the ToMI:SR-Adult represent our initial attempt to address this gap by developing a reliable, valid, and norm-referenced measure of advanced social cognition for adults at risk for poor theory of mind development (norms in the form of percentile rankings by group [ASD/N-ASD] and sex at the composite and item level are available at [theoryofmindinventory.com](http://theoryofmindinventory.com)). A self-report tool is especially important for this population as it provides a vehicle for informants to provide their own insights into their theory of mind understanding or lack thereof.

Given the encouraging psychometric data reported here (reliability, validity, and accuracy of classification), further psychometric evaluation of the ToMI:SR-Adult is justified and is currently underway. Specifically, test–retest reliability will be evaluated to determine the temporal stability of responses and additional measures of convergent validity will be evaluated to determine the degree to which ToMI:SR-Adult scores align (as we expect they will) with measures of autism severity and social-adaptive functioning.

The qualitative data reported here also confirm the value of a self-report ToM tool for people with social learning challenges. Autistic adults in our sample were able to explain their thinking and social cognitive experiences in the context of complex social situations: their accounts were rich and they described the skills they possessed (even if partially), the skills they lacked, and the skills they learned (but that did not come naturally) through a process of reflection, practice, study, therapy, or training. In short, it appears that the autistic adults in our sample were highly aware of, and had much insight into, their own theory of mind functioning.

Nevertheless, there is no contradiction in noting that we also believe that there are some individuals who will lack sufficient awareness of their social cognitive challenges (or an adequate reading comprehension level) and when this is the case, the ToMI:SR-Adult will not be a valid assessment as it relies on the accuracy of respondents’ subjective appraisals. Going forward, an important question will be:

how prevalent is good self-insight in verbal autistic adults? In our study, we solicited responses from an online community who were presumably interested in the notion of social cognition and theory of mind. Because this sample may be unique in their degree of interest, reflection, and self-insight, a broader sample of verbal adults recruited via a different methodology is needed to answer this question. This is also important in light of the comorbidities that often accompany ASD (e.g., other psychiatric conditions, intellectual disability) although our preliminary data are encouraging in this respect as our sample was highly heterogeneous with regard to these demographic variables.

## Clinical Applications

If further validation efforts are successful, the ToMI:SR-Adult can be recommended with confidence as a clinical tool to assess advanced theory of mind skills in adults. Specifically, because it is an informant measure, The ToMI:SR-Adult does not suffer from test-practice effects and may be useful in the development of individualized treatment and progress monitoring. This application is particularly powerful given the wide variability we observed in theory of mind challenges within and between respondents in our sample. As a pre-treatment assessment, the ToMI:SR-Adult can also be used to begin a conversation about treatment priorities and methodologies. Indeed, many of our respondents commented that the survey questions themselves prompted thinking and reflection and, in some cases, revelation and discovery of several theory of mind processes.

Another potential use of the ToMI:SR-Adult is to aid in the screening and identification of ASD. Indeed, data for the measure’s sensitivity and specificity are excellent and the tool could be useful in the screening and identification of ASD among adults who present with high cognitive and language skills and who often go unidentified using existing measures. It is important to reiterate that the ToMI:SR-Adult is intended as a measure of advanced social cognitive skills: it is not a measure of autistic symptoms. As such, the tool may prove useful as part of a larger assessment battery in the identification of other psychiatric, developmental, acquired, and congenital disorders that are also associated with theory of mind impairment (consider sensory loss, attention-deficit hyperactivity disorder, fragile X syndrome, traumatic brain injury).

To the degree that there are broad commonalities in ToM impairments across clinical conditions, the ToMI:SR-Adult is *not* appropriate as a tool for differential diagnosis. On the other hand, some research suggests the presence of syndrome-specific ToM deficiencies and proficiencies (e.g., Buhler et al. 2011; Cornish et al. 2005; Perner et al. 2002). Because the ToMI:SR-Adult taps a wide range of ToM understandings, there is a potential for its use for differential

diagnosis in the future as more information accrues to inform the questions as to whether there are unique ToM profiles associated with specific clinical populations. For the time being, the ToMI:SR-Adult represents a promising research and clinical tool for the assessment of adults with social learning challenges and our data support the contention that many autistic adults possess high degrees of social awareness and are uniquely situated to report on their own ToM experiences.

**Author contributions** TLH, LL, & PAP conceived of and planned the study. TLH and LL carried out the data collection procedures. TLH and AB performed the quantitative analyses. LL performed the qualitative analyses. TLH and LL wrote the manuscript with support by PAP.

## References

- Allman, M., & Falter, C. (2015). Abnormal timing and time perception in autism spectrum disorder? A review of the evidence. In A. Vatakis & M. J. Allman (Eds.), *Time distortions in mind: Temporal processing in clinical populations* (pp. 37–56). Boston: Brill.
- 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). New York: Psychology Press.
- Astington, J. W., & Baird, J. A. (2005). Introduction: Why language matters. In J. W. Astington & J. A. Baird (Eds.), *Why language matters for theory of mind* (pp. 3–25). Oxford: Oxford University Press.
- Baron-Cohen, S. (2009). *The essential difference: Male and female brains and the truth about autism*. New York: Basic Books.
- Baron-Cohen, S., Wheelwright, S., Hill, J., Raste, Y., & Plumb, I. (2001). The “Reading the Mind in the Eyes” Test revised version: A study with normal adults, and adults with Asperger syndrome or high-functioning autism. *The Journal of Child Psychology and Psychiatry and Allied Disciplines*, 42(2), 241–251.
- Bosacki, S. L. (2000). Theory of mind and self-concept in preadolescents: Links with gender and language. *Journal of Educational Psychology*, 92(4), 709.
- Braun, V., & Clarke, V. (2006). Using thematic analysis in psychology. *Qualitative Research in Psychology*, 3(2), 77–101.
- Brien, A., Hutchins, T., & Westby, C. (in press). Autobiographical memory in ASD, ADHD, hearing loss, and childhood trauma: Implications for social communication intervention. *Tutorial: Language, Speech, and Hearing Services in Schools*.
- Buhler, E., Bachman, C., Goyert, H., Heinzl-Gutenbrunner, M., & Kamp-Becker, I. (2011). Differential diagnosis of autism spectrum disorder and attention deficit hyperactivity disorder by means of inhibitory control and theory of mind. *Journal of Autism and Developmental Disorders*, 41, 1718–1726.
- Compton, D., Fuchs, D., Fuchs, L., & Bryant, J. (2006). Selecting at-risk readers in first grade for early intervention. *Journal of Educational Psychology*, 98(2), 394–409.
- Cornish, K., Burack, J. A., Rahman, A., Munir, F., Russo, N., & Grant, C. (2005). Theory of mind deficits in children with Fragile X Syndrome. *Journal of Intellectual Disability Research*, 49(5), 372–378.
- Crehan, E., Altoff, R., Riehl, H., Prelock, P., & Hutchins, T. (2019). Brief report: Me, reporting on myself: Preliminary evaluation of the criterion-related validity of the Theory of Mind Inventory-2 when completed by autistic young adults. *Journal of Autism and Developmental Disorders*. <https://doi.org/10.1007/s10803-019-04278-5>.
- Creswell, J. W., & Plano Clark, V. L. (2017). *Designing and conducting mixed methods research*. Thousand Oaks: Sage.
- Cutting, A. L., & Dunn, J. (1999). Theory of mind, emotion understanding, language, and family background: Individual differences and interrelations. *Child Development*, 70(4), 853–865.
- Davies, M., & Stone, T. (2003). Synthesis: Psychological understanding of social skills. In B. Repacholi & V. Slaughter (Eds.), *Individual differences in theory of mind: Implications for typical and atypical development* (pp. 305–352). New York: Psychology Press.
- Devine, R. T., & Hughes, C. (2013). Silent films and strange stories: Theory of mind, gender, and social experiences in middle childhood. *Child development*, 84(3), 989–1003.
- Dritschel, B., Wisely, M., Goddard, L., Robinson, S., & Howlin, P. (2010). Judgements of self-understanding in adolescents with Asperger syndrome. *Autism*, 14(5), 509–518.
- Happé, F. G. (1994). An advanced test of theory of mind: Understanding of story characters’ thoughts and feelings by able autistic, mentally handicapped, and normal children and adults. *Journal of autism and Developmental disorders*, 24(2), 129–154.
- Hughes, C., Jaffee, S. R., Happé, F., Taylor, A., Caspi, A., & Moffitt, T. E. (2005). Origins of individual differences in theory of mind: From nature to nurture? *Child Development*, 76(2), 356–370.
- Hutchins, T., Bonazinga, L., Prelock, P. A., & Taylor, R. S. (2008). Beyond false beliefs: The development and psychometric evaluation of the perceptions of children’s theory of mind measurement experimental version (PCToMM-E). *Journal of Autism and Developmental Disorders*, 38(1), 143–155.
- Hutchins, T. L., Prelock, P. A., & Bonazinga, L. (2012). Psychometric evaluation of the theory of mind inventory: A study of typically developing children and children with autism spectrum disorders. *Journal of Autism and Developmental Disabilities*, 42(3), 327–341.
- Hutchins, T., 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.
- Kaland, N., Møller-Nielsen, A., Callesen, K., Mortensen, E. L., Gotlieb, D., & Smith, L. (2002). A new advanced’ test of theory of mind: Evidence from children and adolescents with Asperger syndrome. *Journal of Child Psychology and Psychiatry*, 43(4), 517–528.
- Kincaid, J. P., Braby, R., & Mears, J. E. (1988). Electronic authoring and delivery of technical information. *Journal of instructional development*, 11(2), 8–13.
- Klin, A., Schultz, R., & Cohen, D. J. (2000). Theory of mind in action: Developmental perspectives on social neuroscience. In S. Baron-Cohen, H. Tager-Flusberg, & D. J. Cohen (Eds.), *Understanding other minds: Perspectives from developmental neuroscience* (pp. 357–390). Oxford, UK: Oxford University Press.
- Lai, M.-C., & Baron-Cohen, S. (2015). Identifying the lost generation of adults with autism spectrum conditions. *The Lancet Psychiatry*, 2(11), 1013–1027. [https://doi.org/10.1016/S2215-0366\(15\)00277-1](https://doi.org/10.1016/S2215-0366(15)00277-1).
- Lerner, M., Hutchins, T., & Prelock, P. (2011). Brief report: Preliminary evaluation of the theory of mind inventory and its relationship to measures of social skills. *Journal of Autism and Developmental Disorders*, 41, 512–517.
- Lewis, L. F. (2017). A mixed methods study of barriers to formal diagnosis of autism spectrum disorder in adults. *Journal of Autism*

- and *Developmental Disorders*, 47(8), 2410–2424. <https://doi.org/10.1007/s10803-017-3168-3>.
- McCauley, R. (2001). *Assessment of language disorders in children*. Mahwah, NJ: Lawrence Erlbaum.
- Mosner, M., Kinard, J., Shah, J., McWeeny, S., Greene, R., ... Dichter, G. (2019). Rates of co-occurring psychiatric disorders in autism spectrum disorder using the Mini International Neuropsychiatric Interview. *Journal of Autism and Developmental Disorders*, 49, 3819–3828.
- 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.
- Nelson, K. (2003). Narrative and self, myth and memory: Emergence of the cultural self. In R. Fivush & C. A. Haden (Eds.), *Autobiographical memory and the construction of a narrative self: Developmental and cultural perspectives* (pp. 3–28). Mahwah: Lawrence Erlbaum Associates Publishers.
- Park, S., Goo, J., & Jo, C. (2004). Receiver operating characteristics (ROC) curve: Practical review for radiologists. *Korean Journal of Radiology*, 5(1), 11–18.
- Pears, K. C., & Moses, L. J. (2003). Demographics, parenting, and theory of mind in preschool children. *Social Development*, 12(1), 1–20.
- Perner, J., Kain, W., & Barchfeld, P. (2002). Executive control and higher-order theory of mind in children at risk for ADHD. *Infant and Child Development*, 11, 141–158.
- Rosen, R., Mazefsky, C., Vasa, R., & Lerner, M. (2018). Co-occurring psychiatric conditions in autism spectrum disorder. *International Review of Psychiatry*, 30(1), 40–61.
- Schriber, R., Robins, R., & Solomon, M. (2014). Personality and self-insight in individuals with autism spectrum disorder. *Journal of Personality and Social Psychology*, 106(1), 112–130.
- Slaughter, V., & Repacholi, B. (2003). Introduction: Individual differences in theory of mind. In B. Repacholi & V. Slaughter (Eds.), *Individual differences in theory of mind* (pp. 1–12). New York: Taylor & Francis Books.
- Spreng, R. N., & Andrews-Hanna, J. R. (2015). The default network and social cognition.
- Spreng, R. N., Mar, R. A., & Kim, A. S. N. (2008). The common neural basis of autobiographical memory, prospection, navigation, theory of mind, and the default mode: A quantitative meta-analysis. *Journal of Cognitive Neuroscience*, 21, 489–510. <https://doi.org/10.1162/jocn.2008.21029>.
- Yuhas, D. (2018). Ask me first: What self-assessments can tell us about autism. *Spectrum: Autism Research News*. [www.spectrumnews.org](http://www.spectrumnews.org)
- White, S., Hill, E., Happé, F., & Frith, U. (2009). Revisiting the strange stories: Revealing mentalizing impairments in autism. *Child Development*, 80(4), 1097–1117.

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