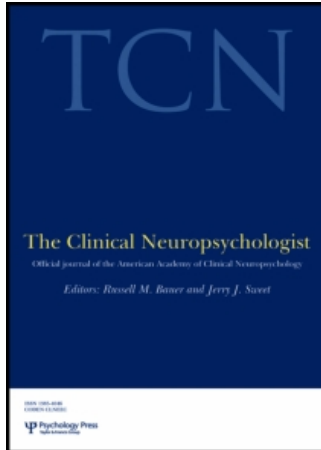


This article was downloaded by:[Canadian Research Knowledge Network]  
On: 21 August 2007  
Access Details: [subscription number 770885180]  
Publisher: Psychology Press  
Informa Ltd Registered in England and Wales Registered Number: 1072954  
Registered office: Mortimer House, 37-41 Mortimer Street, London W1T 3JH, UK



## The Clinical Neuropsychologist

Publication details, including instructions for authors and subscription information:  
<http://www.informaworld.com/smpp/title~content=t713721659>

### Developing Clinically Suitable Measures of Social Cognition for Children: Initial Findings from a Normative Sample

Online Publication Date: 01 March 2007

To cite this Article: Saltzman-Benaiah, Jennifer and Lalonde, Christopher E. (2007) 'Developing Clinically Suitable Measures of Social Cognition for Children: Initial Findings from a Normative Sample', *The Clinical Neuropsychologist*, 21:2, 294 - 317  
To link to this article: DOI: 10.1080/13854040500473760  
URL: <http://dx.doi.org/10.1080/13854040500473760>

PLEASE SCROLL DOWN FOR ARTICLE

Full terms and conditions of use: <http://www.informaworld.com/terms-and-conditions-of-access.pdf>

This article maybe used for research, teaching and private study purposes. Any substantial or systematic reproduction, re-distribution, re-selling, loan or sub-licensing, systematic supply or distribution in any form to anyone is expressly forbidden.

The publisher does not give any warranty express or implied or make any representation that the contents will be complete or accurate or up to date. The accuracy of any instructions, formulae and drug doses should be independently verified with primary sources. The publisher shall not be liable for any loss, actions, claims, proceedings, demand or costs or damages whatsoever or howsoever caused arising directly or indirectly in connection with or arising out of the use of this material.

© Taylor and Francis 2007

## DEVELOPING CLINICALLY SUITABLE MEASURES OF SOCIAL COGNITION FOR CHILDREN: INITIAL FINDINGS FROM A NORMATIVE SAMPLE

Jennifer Saltzman-Benaiah and Christopher E. Lalonde

*Department of Psychology, University of Victoria, Victoria, B.C., Canada*

*Our understanding of children's social competence has increased tremendously over the past two decades. There is increasing evidence to suggest that social-cognitive impairments are not restricted to children on the autistic spectrum, but rather may be associated with a host of developmental and acquired neurological conditions including learning disabilities, attention deficit disorder, traumatic brain injury, and stroke. Although many investigators have begun to bridge the gap between clinical practice and research by applying experimental tasks to clinical populations, few tools are available for the clinical evaluation of social competence, particularly in children. This study marks a series of first steps in the development of measures suitable for the assessment of children between 6 and 12 years of age. The results of the study provide data for a number of experimental tasks that have been adapted with clinical practice in mind. A discussion of the developmental progressions and the relationships among the measures is also included.*

### INTRODUCTION

Social interaction represents a significant part of our daily activity. We often take for granted the skills that are involved in effective social interaction because, in most people, these skills develop naturally without specific attention or teaching. Knowing what is funny, what is rude, and when “enough is enough” are examples of the subtle aspects of social interaction that seem to rely on gut feelings or social intuitions, rather than concrete or rule-based decision-making processes. Effective social interaction requires a complex blend of knowledge, interpretation, inhibition, and action. Since knowledge and interpretation often guide our actions, it is not surprising that misinterpretation is frequently at the root of inappropriate behavior. Unfortunately, inappropriate social behavior is often met with embarrassment or

---

Address correspondence to: Jennifer Saltzman-Benaiah, Ph.D., C.Psych., Tourette Syndrome Neurodevelopmental Clinic, Toronto Western Hospital, 3rd Floor West Wing, 399 Bathurst Street, Toronto, Ontario, Canada M5T 2S8. Tel.: 416-603-5794. Fax: 416-603-5180. E-mail: jsaltzma@uhnres.utoronto.ca

Jennifer Saltzman-Benaiah is now at the Toronto Western Hospital, University Health Network, Toronto, Ontario.

The first author received funding support from a University of Victoria Graduate Fellowship during the completion of this project.

Accepted for publication: October 26, 2005. First published online: July 22, 2006.

scorn, resulting in punishment for an action that may not have been intentional. A more complete understanding of the cognitive contributions to social behavior is critical if we are to accurately assess and manage inappropriate social behavior or remediate behavior problems.

A role for neuropsychology in understanding, detecting, and treating cognitive aspects of social behavior seems natural, since clinical neuropsychologists are often called upon to provide comprehensive assessments of patients whose cognitive and behavioral difficulties may be related to underlying neurological dysfunction. However, the lack of standardized measures of social cognitive skills may restrict the clinical neuropsychologist's evaluation of social behavior to third-party reports of a patient's difficulties. Fortunately, a growing base of knowledge from the child development and cognitive neuropsychology literatures may provide important clues as to some of the skills that are associated with effective social interaction and introduce more objective means of evaluating social cognitive skills in clinically referred children.

The number of arguably separable content areas beneath the umbrella of "social cognition" is vast, and too extensive to be adequately reviewed within the scope of this article. We have chosen instead to concentrate our attention on a small set of areas that show particular promise for the development of clinical tools. Among the most widely researched and familiar aspects of social cognition in recent years is the concept of "theory of mind," sometimes referred to as "mentalizing" abilities (Terwogt & Rieffe, 2003; Astington, Harris, & Olsen, 1988; Wimmer & Perner, 1983). Having captured the interest of both developmental and cognitive neuropsychological researchers, it seems a reasonable place to begin looking for clinically suitable measures that could supplement an assessment carried out within a developmental neuropsychological framework. Other aspects of children's social cognitive development (such as display rules) have also been documented using various tasks and may be adaptable for clinical use.

## FALSE BELIEF AND INTENTIONALITY

One of the more widely studied milestones of social cognitive developments is the acquisition of an understanding of false belief. That is, the understanding that mental states do not always reflect reality and that people can hold (and act upon) beliefs that are objectively false. A common false belief scenario involves the recognition that a person may hold an incorrect belief because he or she was not privy to a relevant piece of information. For example, Johnny *thinks* that the scissors are in the drawer, because he did not *see* his sister take them. Results from two decades' worth of research using a roster of false belief measures have documented a clear turning point in the development of false belief understanding during the preschool years: 3-year-olds typically fail such tasks, while 4-year-olds just as typically succeed.

Children who understand that beliefs can be false quickly learn that false beliefs can be engineered; in other words, they learn how to deceive and how to lie (Sodian, 1991). Once children grasp the relation between perceptual access, belief formation, and behavior, they also seem to realize that by manipulating the perceptions and experiences of others, one can also manipulate the types of beliefs they might hold (Hala, Chandler, & Fritz, 1991; O'Neill & Gopnik, 1991).

In addition to deliberate lies, other forms of communication—such as jokes and sarcastic comments—also rely upon an awareness of mental states and intentions

(Leekham, 1991; Perner, 1988). A sarcastic remark, for example, is meant to contain both a true and false meaning—a compliment and a criticism (e.g., “Your new shoes look *great*”). Recognizing the double-edged *intention* in such communications appears to be more difficult for children (and often for adults) than attributing a simple false belief based on ignorance or lack of perceptual access to the truth. Leekham (1991) reports that it is not until the age of 8 or 9 years that children are consistently able to distinguish a joke from a deceitful remark. Similarly, Demorest, Meyer, Phelps, Gardner, and Winner (1984) noted improvements in the identification of deliberate lying between 6 and 13 years of age, and improvement in identification of sarcastic remarks between 13 years and adulthood. Such findings suggest that acquiring a theory of mind (i.e., understanding what others “know”) demands more than a simple appreciation of what they see or what they say.

### INTERPRETIVE THEORY OF MIND: BEYOND FALSE BELIEF

Although false beliefs are sometimes related to perceptual access, this alone cannot explain how it is that, given exactly the same information, two people can still come to hold different beliefs. In response to the ongoing debate over theory of mind development beyond the preschool years, several researchers have proposed the existence of a more subtle and complex level of mental state understanding. According to one explanation, children’s ongoing development in middle childhood represents a transition from holding a *copy* theory of mind to an *interpretive* theory of mind (Chandler, 1988; Chandler & Helm, 1984). Individuals in possession of a “copy” theory of mind are able to understand that what a person sees can determine what they know—in other words, they appreciate that direct perception can lead to knowledge—beliefs are a copy of reality. In contrast, those with an “interpretive” theory of mind, understand that, even when granted unfettered access to all of the essential information about some event, people are still free to arrive at different beliefs about the same event. Consider, for example, Margaret and Tim, who are looking at a piece of abstract art. Margaret *thinks* it is a picture of a dog, while Tim *thinks* it is a flower. Copy theories cannot account for these different beliefs.

Some have proposed that it is not until middle childhood that children begin to appreciate that a single situation, object, or event—even witnessed in exactly the same way by two individuals—can still be subject to multiple interpretations (Chandler & Lalonde, 1996; Carpendale & Chandler, 1996; Lalonde & Chandler, 2002). Carpendale and Chandler (1996) argue that “interpretation is a thing needing to be done . . . in response to ambiguities brought on by a shortage of relevant information” (p. 1693). By these lights, children are said to possess an interpretive theory of mind only when they are able to appreciate the interpretive possibilities inherent in ambiguous situations. Passing a standard false belief task is no guarantee of success in such circumstances and it is not until the age of 7 or 8 years old that children consistently are able to recognize that contrasting interpretations might be the result of an ambiguous situation (Carpendale & Chandler, 1996; Lalonde & Chandler, 1995, 2002).

### STRATEGIC CONTROL OF EMOTIONS: USING DISPLAY RULES

While success on false belief and deception tasks appears to reflect an understanding that beliefs need not always accord with reality, a growing number

of investigators have examined the use of *display rules* as a means to infer whether or not children understand that expressed emotions are not always the same as experienced emotions (Zeman & Garber, 1996; Gnepp & Hess, 1986; Terwogt & Olthof, 1989). A display rule has been defined as “a principle guiding when and how people regulate their emotional expressions” (Gnepp & Hess, 1986).

Studies of children’s understanding of display rules have included observational studies of children attempting to hide their emotions in disappointing situations (Saarni, 1984), as well as children’s self reports concerning the actions they would take in order to hide their emotions in different situations (Zeman & Garber, 1996; Harris, Donnelly, Guz, & Pitt-Watson, 1986; Gnepp & Hess, 1986). Gnepp and Hess (1986) found that children’s understanding of display rules increased steadily between grades 1 through 5, and that children were more likely to express display rules in situations of prosocial rather than self-protective motivation. The authors also reported that children’s understanding of *verbal* display rules (i.e., saying something to hide your feelings) outpaced their understanding of *facial* display rules (i.e., changing your outward expression to hide your feelings). Other studies have generally supported the finding that school-aged children develop an awareness of different ways to control overt emotional expressions.

By acknowledging and demonstrating the use of display rules to hide a true emotional expression, children appear to recognize that emotions, like beliefs, can be intentionally falsified, and that another person’s perception of one’s emotional state—and thereby their beliefs—can be influenced by intentionally deceptive means (i.e., the strategic use of display rules).

## PERSONALIZED INFERENCES IN COGNITION AND EMOTION

In developing an interpretive theory of mind, children come to understand that knowledge is constructed by the perceiver, rather than simply copied from perceptual experiences. Understanding that people make contributions to the way information is perceived and beliefs are formed, underpins the appreciation that the same information may be interpreted in different ways. A further development in children’s understanding is the appreciation that differences in background knowledge, personal traits, or any number of individual differences can contribute to the diversity of interpretations that can arise among participants in the same situation. Studies have shown that young children who pass standard false belief tasks are not always able to recognize that individual differences in background knowledge can result in different interpretations (Lalonde & Chandler, 2002; Carpendale & Chandler, 1996). In particular, researchers have examined children’s ability to use personal information about a character’s age (baby vs. adult) in making judgments about the character’s knowledge in different situations (Montgomery, 1993; Taylor, Cartwright, & Bowden, 1991). Montgomery (1993) found that children between the ages of 6 and 9 years old tend to overestimate a baby’s ability to benefit from verbal communication regarding the location of a hidden object. Children tended to report that, like adults, babies could learn about the location of a hidden object if they were given verbal information. Only 33% of eight- to nine-year-olds were able to understand: (a) that ambiguous statements do not convey information about an object’s location, and

(b) that a baby's lack of language skill would preclude the understanding of any statement, ambiguous or otherwise, about an object's location.

Similarly, Gnepp and Gould (1985) examined adults' and children's ability to use information about an individual's prior experiences to predict that person's emotional response in a particular situation. Participants heard a story describing a character who was involved in two separate events occurring one after another. The events were presented such that one might reasonably expect the experience of the first event to influence the character's feelings towards the second event. According to the concepts of *personalized* and *situational* responses described by Gnepp and Gould (1985), a personalized response about the character's emotional reaction to the second event would take into account the influence of the first event. On the other hand, a situational response would ignore the influence of the first event, and focus entirely on the second. Children's ability to spontaneously determine the relevance of previous events and apply this knowledge in making judgments developed steadily between 1st grade and college years. Only college students made reliable and consistent use of this information, while 5th grade students failed to make personalized responses about one-third of the time.

### **IMPAIRMENTS IN SOCIAL COMPETENCE: EMPIRICAL KNOWLEDGE MEETS CLINICAL UTILITY**

Although investigations into children's developing understanding of beliefs and emotions have traditionally focused on typically developing children, an increasing number of studies of children (and adults) with apparent deficits are now available.

Within the theory of mind literature, there are many reports suggesting that individuals with autistic spectrum disorders have considerable difficulty on theory of mind tasks relative to their non-autistic age peers (e.g., Baron-Cohen, Tager-Flusberg, & Cohen, 2000; Leekham & Perner, 1991; Happé, 1993). While some higher-functioning individuals with autism are able to successfully complete first-order false belief tasks, they often fail on so-called second-order tasks, in particular, those involving metaphor, simile and irony (Happé, 1993). The findings in this area have been so numerous that some have proposed that failure to develop mental state understanding is *the* principal deficit in autism (see Rogers & Pennington, 1991, for a discussion), and researchers have begun to explore underlying brain mechanisms that account for this core deficit (Rogers & Pennington, 1991; Brothers, 1990).

Although reports of theory of mind performance among children with other disabilities are few, investigations of other related difficulties in everyday behaviors such as making conversation, reading emotional and social cues, and controlling emotional expression have been reported in children with attention deficit disorder (ADHD; Matthys, Cuperus, & van Engeland, 1999; Milch-Reich, Campbell, Pelham, Connelly, & Geva, 1999; Greene et al., 1996), learning disabilities (Sprouse, Hall, Webster, & Bolen, 1998; Rourke & Tsatsanis, 1996), mental retardation (Moffatt, Hanley-Maxwell, & Donnellan, 1995; Adams & Markham, 1991), and head injury (Dennis, Barnes, Wilkinson, & Humphreys, 1998; Petterson, 1991; Jackson & Moffatt, 1987).

In the neuropsychological literature, studies investigating the neural bases of social cognition (primarily theory of mind) have increased dramatically in the past

10–15 years. Initial studies reported correlations with executive functioning, thereby positing a role for the frontal lobes in theory of mind (McEvoy, Rogers, & Pennington, 1993; Ozonoff, Pennington, & Rogers, 1991; Zelazo & Mueller, 2002). Studies of individuals with focal frontal lesions or degenerative disease processes, such as frontotemporal dementia, have lent further support to this claim (Gregory et al., 2002; Snowden et al., 2003; Stone, Baron-Cohen, & Knight, 1998). Most recently, neuroimaging technology has been engaged to address the question of which brain regions are uniquely involved in theory of mind functions. With a high rate of consistency, studies have demonstrated activation in specific brain regions, including the orbitofrontal cortex, paracingulate cortex, and superior temporal sulcus (e.g., Baron-Cohen et al., 1994; Grezes, Frith, & Passingham, 2004; Gallagher & Frith, 2003). As a result, neuroscientists have begun to postulate theories of social cognitive brain functioning including, but not limited to, theory of mind and conscious awareness (e.g., Stuss & Anderson, 2004).

While the findings of these studies certainly add to our knowledge of social competence in various disorders (and the underlying brain mechanisms that may be responsible), there are some limitations from the standpoint of clinical utility. First, each study has been conducted independently by separate researchers, using different tasks and different methodologies. Because of this, it is difficult to relate the nature and extent of social impairments across clinical groups. Is the type of difficulty experienced by one group (e.g., ADHD) qualitatively and/or quantitatively different from another group (e.g., nonverbal learning disability)? Are all researchers describing the same impairment when they make reference to “social disability” or the plethora of other commonly used terms?

Second, although the tasks employed by experimental designs may be appropriate for group comparison, they are not always appropriate for clinical use. For the most part, experimental designs demonstrate statistical differences by comparing the performance of a clinical and non-clinical group on the same measure. Notwithstanding the fact that statistical significance does not always equate with clinical impairment, the results derived from an experimental format provide little information about the degree of impairment of an *individual* child relative to his/her peers. These studies are designed to look at overall group differences without providing interpretation at the individual level. In many studies, individual subject performance on measures of social competence is highly variable. Finally, many of the protocols are lengthy, use dichotomous (pass-fail) rather than continuous scoring systems, or are based on studies with a restricted age group (e.g., only 3-year-olds).

The current study represents an attempt to move beyond these limitations toward a clinically appropriate tool. In order to begin to develop tasks suitable for evaluating social interaction skills in school-aged children, there were several important factors to consider. To be useful to clinicians, the measure should: (a) assess a broad range of skills and abilities related to social interaction, (b) provide normative data across a wide range of ages to allow for single case (clinical) use and interpretation, (c) have a relatively short administration time to facilitate integration within larger testing batteries, and (d) demonstrate ecological validity and relevance to everyday behaviour and clinical practice.

By selecting tasks in five content areas: (a) false belief, (b) interpretive theory of mind, (c) display rules, (d) making personalized inferences about thoughts, and (e) making personalized inferences about emotions, we set out to provide a preliminary base of

normative information while investigating the relationship between tasks and everyday behavior, with an eye to future studies examining their utility in clinical populations.

## METHOD

### Participants

A total of 89 children (38 males) were recruited across 7 age groups (6-year-olds, 7-year-olds, 8-year-olds, 9-year-olds, 10-year-olds, 11-year-olds, 12-year-olds). Although SES data was not collected on individual children, participating schools were located in upper-middle class neighbourhoods in the Greater Toronto area. Children with reported developmental, learning and/or attentional difficulties (e.g., ADHD, NVLD, PDD) were excluded from the study, as were participants with any history of neurological disorder or acquired brain injury. During recruitment, school principals and teachers were asked to identify students in their class who had known diagnoses as well as students who, despite not having a specific diagnosis, had been formally identified as a students with special learning needs. This information was further confirmed by a short background history questionnaire that was completed by parents of study participants.

The single inclusion criterion was fluency in English. All subjects received the Vocabulary subtest of the *Wechsler Intelligence Scale for Children*, third edition (WISC-III; Wechsler, 1991). Scores on this measure were used to confirm that all subjects fell broadly within the Average range of intellectual functioning (i.e., within 1 standard deviation). None of the participants tested scored below the average range, and therefore no one was excluded from the study on these grounds. A summary of the sample demographics is presented in Table 1.

### Procedure

Children were recruited to participate in a protocol that was approved by the University of Victoria Ethics Board and the York Region Board of Education (in a suburb of Toronto) by means of an information letter sent home to parents by the child's school. Parents who agreed to have their child participate in the study were invited to complete and return a consent form and the Parent Questionnaire to their

**Table 1** Demographics of the sample

Age group (Year olds)	N	Sex		Age Mean (SD)	WISC-III Vocabulary SS	
		M	F		Mean (SD)	Range
6	12	5	7	6.6 (0.2)	11.3 (2.1)	8–15
7	14	7	7	7.6 (0.3)	11.9 (1.9)	9–16
8	10	4	6	8.6 (0.3)	11.5 (1.9)	10–15
9	17	10	7	9.6 (0.2)	12.4 (1.5)	10–16
10	14	3	11	10.4 (0.3)	10.6 (2.1)	7–14
11	11	4	7	11.5 (0.4)	12.0 (1.5)	9–15
12	11	5	6	12.6 (0.3)	11.3 (1.8)	8–14
Total	89	38	51	9.5 (1.9)	11.6 (1.9)	7–16

child's teacher. Teacher questionnaires were then distributed for children who had been given consent to participate. Each child participated in a one-on-one session with the first author (JSB) at his or her school during regular school hours. After the purpose of the study was explained, each child was asked to give his or her assent to the procedure. Upon completion of their sessions, children were rewarded for their participation with a colourful certificate bearing their name and the signature of the experimenter.

### **Measures**

Each of the five measures developed for this study was based on tasks and procedures described in previous studies. For the purposes of this study, the measures were entitled: (a) False Belief, Intentional Deception & Sarcasm, (b) Interpretive Theory of Mind, (c) Display Rules, (d) Personalized Thoughts, and (e) Personalized Emotions. In selecting appropriate tasks, important considerations included: (a) whether the same measure could be used across the entire age span, (b) whether the measure could yield continuous data (i.e., was not pass/fail), and (c) task administration time. Although the formats for the tasks in this study were based on methods reported in previous research, all of the stories and pictures were created uniquely for this study.

All stories and instructions were read aloud and repeated as many times as required in order to ensure comprehension. Stories were accompanied by a picture (depicting the general scene and characters) that remained in front of the child throughout the duration of the story and the questions. All participants were tested by a single trained examiner who attempted to read the stories consistently. Administration of the entire protocol (including screening measures) took approximately 30 to 50 minutes depending on the age of the child. Table 2 provides a brief description of each task including administration and scoring, and a reference to the original study on which the task was based. Sample stories are included in Appendix A.

### **Parent and Teacher Questionnaires**

The questionnaires were designed to elicit parent and teacher reports of some everyday behaviors that might reasonably involve an understanding of mental states and emotions, including social interaction scenarios that would be commonly experienced by a school-aged child (e.g., making friends, being teased, understanding a joke), as well as specific skills or behaviors that have been reported empirically as areas of weakness in certain groups of children (e.g., autism).

Designed to be accessible to parents with minimal reading levels, the Parent's Questionnaire consisted of 57 statements that the parents rated on a 4-point Likert-type scale (0 = never, 1 = sometimes, 2 = often, 3 = almost always). The Teacher's Questionnaire represented an abbreviated (33-item) version of the parent's questionnaire, including the items deemed most applicable to the school environment. Sample items are included in Appendix B.

### **Reliability**

All scores requiring judgment of verbal responses were scored independently by two examiners to determine interrater reliability. Also, in order to examine the

**Table 2** Brief task descriptions, administration and scoring guidelines

Task	Description	Administration	Scoring
False Belief, Intentional Deception and Sarcasm (e.g., Demorest et al., 1984)	8 short stories depicting interpersonal situations involving false belief, deception or sarcasm	Each story is read accompanied by an illustration of the main characters interacting. Children are asked to make a <i>prediction</i> and provide an <i>explanation</i> about one of the character's beliefs in that situation.	One point per story for each correct prediction; Total = 8 points
Interpretive Theory of Mind (Carpendale & Chandler, 1996)	3 short vignettes in which 2 characters (puppets) express different interpretations of the same (but ambiguous) stimuli. Stimuli included represented situations of pictorial ambiguity (e.g., a visual illusion), referential ambiguity (e.g., the red one), and lexical ambiguity (e.g., Get me a pair/pear).	Each vignette is enacted using the puppets while the child watches. Children answer 3 questions per vignette: (1) Is it okay for the 2 characters to hold different views of the stimulus? (2) Can we predict how another person would interpret the ambiguous stimuli? (3) Is every interpretation equally likely and/or valid?	One point per question correctly answered; Total = 9 points
Display Rules (Gnepp & Hess, 1986)	6 brief stories about children in situations in which people would want to hide their true emotions. 3 stories are designed to elicit prosocially motivated display rules (protecting other people) and 3 motivated by self-protection (e.g., from embarrassment).	After hearing each story, children are asked to decide what the character in each story would say, and what kind of face they would make.	Children's responses are coded as "display rule" (1 point) or "other" response (0 points); Total = 12 points

Personalized Thoughts (Montgomery, 1993)	Four brief scenes are enacted using dolls and props. In each scene a "child" doll gives information about the location of a hidden prize to another doll (either adult or infant). The information is either ambiguous or unambiguous.	Each scene is enacted using two dolls. Children are asked if the adult (or infant) doll knows where the prize is hidden based on the information given by the child doll.	The adult will benefit from unambiguous information, while the infant will not. 1 point is assigned for each correct answer. Total = 4 points
Personalized Emotions (Gnepp & Gould, 1985)	3 short stories describing one character who was involved in 2 separate events. In each case, the stories involve situations in which the first event might influence the character's feelings toward the second event (e.g., seeing a scary movie about snakes and then being scared of a snake).	After hearing the story, children are asked to predict and explain the character's feelings during the second event (e.g., How did the character feel when he saw the snake? Why?)	Responses are scored as "personalized" (1 point) if the character's emotion reflected the influence of the first event. Total = 3 points

test-retest reliability of these measures, a smaller sample of participants was selected to complete the protocol a second time. This sample included two representatives from each age group ( $N = 14$ ). The number of days between the first and second administration ranged from 18 to 35, with a mean difference of 25 days ( $SD = 7$ ).

## RESULTS

For the purpose of analysis, all raw scores were converted into percent correct based on the total number of items in the set. A summary of the results for each of the tasks by age is shown in Figure 1. In order to balance excessive type 1 errors while maintaining reasonable power, the alpha level for post hoc comparisons among subsets of the tasks was set at .01.

### Task 1: False Belief, Intentional Deception and Sarcasm Stories

These data were analyzed using a 2 (gender) by 7 (age) by 3 (story-type: False Belief, Intentional Deception, Sarcasm) mixed model ANOVA with gender and age as between subjects factors and story-type as a within subjects factor. The linear and quadratic effects of age were of particular interest, and were evaluated regardless of the omnibus age effects.

The significant main effect of gender ( $F(1,75) = 5.19, p < .03$ ) indicated that, on average, girls ( $M = 67.31$ ) tended to outperform boys ( $M = 58.43$ ). There were

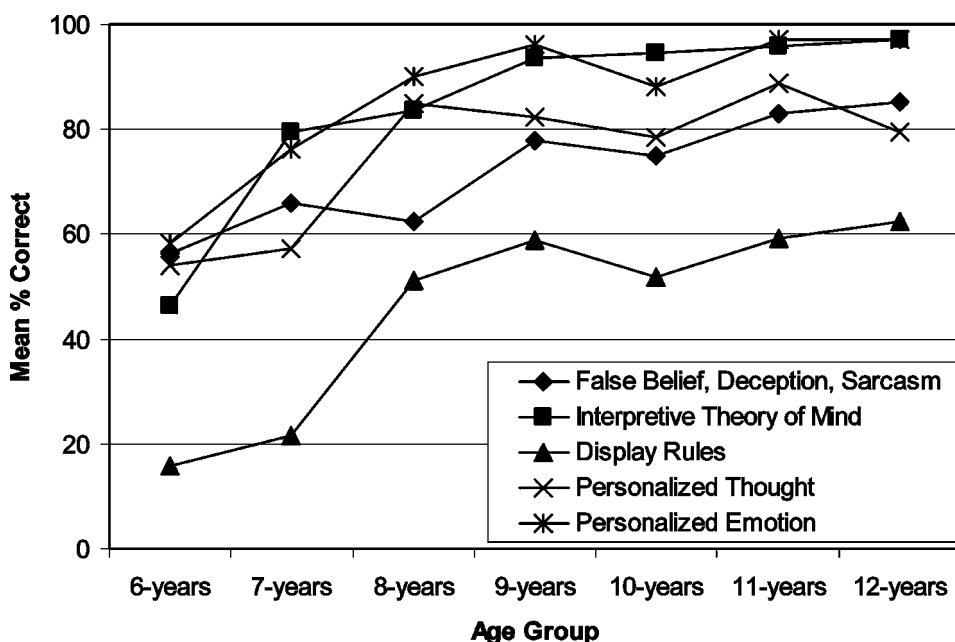


Figure 1 Mean percentage of correct responses on each task by age group.

also significant main effects of age ( $F(6,75) = 6.05, p < .001$ ), and story-type ( $F(2,150) = 132.25, p < .001$ ). The overall age trend was linear ( $p < .001$ ).

Further analyses (repeated measures ANOVA for each paired combination of story types) revealed significant differences between the participants' performance on each of the measures. False belief story performance was significantly better than both intentional deception ( $F(1,75) = 23.53, p < .001$ ) and sarcasm ( $F(1,75) = 243.59, p < .001$ ). Performance on the intentional deception stories was significantly higher than performance on stories involving sarcasm ( $F(1,75) = 110.31, p < .001$ ).

### **Task 2: Interpretive Theory of Mind**

These data were analyzed using a 2 (gender) by 7 (age) by 3 (interpretive task: lexical, pictorial, and referential interpretation) mixed model ANOVA with gender and age as between subjects factors and interpretive task as a within subjects factor. The linear and quadratic effects of age were of particular interest, and were evaluated regardless of the omnibus age effects. The main effect of gender approached significance ( $F(1,75) = 3.41, p < .07$ ), suggesting that, on average, girls' performance ( $M = 87.34$ ) was somewhat better than boys' ( $M = 80.75$ ). There were also significant main effects of age ( $F(6,75) = 16.33, p < .001$ ), and Interpretive task ( $F(2,150) = 10.76, p < .001$ ). The overall age effect was quadratic ( $p < .001$ ). The omnibus interaction between age and interpretive task was not significant. It appears that the greatest improvements in scores may occur between the ages of 6 and 8 years, with a flattening of scores as they reach a ceiling for children between 9 and 12 years old. Participants' performance on the lexical and pictorial interpretation conditions were significantly different from each other ( $F(1,82) = 17.658, p < .001$ ). These data (although cross sectional) appear to suggest that children begin at 6 years of age with a better grasp of pictorial than lexical interpretation, but seem to achieve an equal understanding of each condition by 10 to 12 years of age. This observation is supported by analyses that reveal an interaction between the lexical-pictorial difference and age group that approached significance ( $F(6,82) = 2.66, p < .03$ ). Similarly, the difference between the lexical and referential interpretation conditions was significant ( $F(1,82) = 19.481, p < .001$ ), where performance on the lexical and referential conditions appeared to develop at a comparable rate with children's performance on the referential condition being superior at any given age. In support of this, there was no interaction between the lexical-referential difference score and age group. Finally, there was no significant difference between the referential and pictorial interpretation conditions, nor was there an interaction of this difference with age.

Overall, these results suggest that an understanding of the concepts of referential and pictorial interpretation begins to develop at an earlier age than lexical interpretation, although all three concepts seem to be generally well understood by about 10 to 12 years of age.

### **Task 3: Display Rules**

The data were analyzed using a 2 (gender) by 7 (age) by 2 (display rule type: verbal and facial) mixed model ANOVA with gender and age as between subjects factors and display rule type as a within-subjects factor. The main effect

of gender was not significant. However, there were significant main effects of age ( $F(6,75) = 8.47, p < .001$ ) and display rule type ( $F(1,75) = 48.60, p < .001$ ). The overall age effect was quadratic ( $p < .02$ ). The omnibus interaction between age and display rule type was also significant ( $F(6,75) = 2.60, p < .03$ ). Thus, separate analyses focusing on the linear and quadratic age trends for each display rule type were run.

From visual inspection of the data, it appears that the most dramatic increases in children's performance are occurring between 6 and 8 years of age. In this range, children improved from achieving an average of 15% correct at 6 years of age, to an average of 51% correct at 8 years of age. While this bears some resemblance to the children's pattern of performance on the interpretive measures, scores for display rules do not reach ceiling, but flatten out at about 60% by age 9.

#### **Task 4: Personalized Thoughts**

Overall, participants' performance on this task was significantly related to age ( $r = .378, p < .001$ ) but not gender. The pattern of development appeared again to be quadratic ( $F = 6.075, p = .016$ ). The question was then asked: Does the participants' performance reflect merely a difficulty with the issue of ambiguity, or did the children fail to understand the difference between giving instruction to an adult or an infant? The results of paired samples *t*-tests showed a significant difference between children's performance on the mother and infant conditions ( $t = 8.584, p < .001$ ), but no difference in performance on the ambiguous versus unambiguous conditions. Children performed more poorly on the conditions involving the infant rather than the adult. The fact that they did not perform differently on the infant-ambiguous versus infant-unambiguous conditions suggests that the reason for failure on the infant conditions was not related to the understanding of ambiguity. In fact, children often tried to use the same reasoning in the infant and mother conditions (i.e., like an adult, an infant can benefit from an unambiguous instruction, but not an ambiguous one), and thus failed to recognize the important difference in dealing with adults versus infants.

#### **Task 5: Personalized Emotions**

An analysis of children's performance on the measure of personalized emotions, a two-way ANOVA revealed a significant main effect of age group ( $F = 6.017, p < .001$ ) but not gender. Further analyses revealed a significant quadratic trend in children's performance by age ( $F = 6.806, p = .011$ ). The correlation with age was highly significant ( $r = .473, p < .001$ ).

Once again, dramatic increases in performance were noted between 6 and 8 years of age, with a relative flattening of scores between 9 and 12 years of age. Children in the youngest age group (6 years old) demonstrated an average performance of 58% on this measure, while children between the ages of 9 and 12 years achieved average scores ranging from 88 to 97%, suggesting a possible ceiling effect in their performance.

#### **Reliability**

Inter rater reliability for the measures ranged from .83 to .93, with the exception of Personalized Emotion, which was .64. Test-retest reliability for the measures

**Table 3** Correlations among the measures

Measure	1	2	3	4	5	6	7
1. False-Belief, Deception & Sarcasm	—	.518**	.540**	.355**	.462**	.307**	.255*
2. Interpretive Theory of Mind		—	.391**	.338**	.461**	.288**	.262*
3. Display Rules			—	.472**	.444**	.337**	.165
4. Personalized Thought				—	.291*	.113	−0.051
5. Personalized Emotion					—	.203	.219*
6. Parent Questionnaire						—	.248*
7. Teacher Questionnaire							—

\* $p < .05$ , \*\* $p < .01$ .

was quite variable, ranging from .28 (False Belief) to .93 (Display Rules). Reliability for Interpretive TOM and Personalized Emotion were also reasonably high ( $r = .82$  and  $r = .84$  respectively) while the reliability of the Personalized Thought task was relatively weaker over time ( $r = .45$ ).

### Relationship among Measures: Correlations and Developmental Trends

The interrelationships of scores for the tasks and questionnaires are presented in Table 3. Although it was not within the scope of this study to examine the nature of any causal relationships in the development of these measures, it was of interest to evaluate whether the skills required for these tasks appeared to develop in a particular order. In order to examine this, the following question was asked: At what age do at least 50% of children achieve a perfect score on each task? The results are summarized in Table 4, illustrating the age at which each of the tasks is first passed by a majority of children.

### Parent and Teacher Questionnaires—Content and Ecological Validity

Parent and Teacher Questionnaires were completed for each child who participated in the study. At the time of the study, the teachers had known their students for approximately 6 to 7 months. Analyses were conducted to examine elements of each questionnaire individually, as well as the relationship between parent and teacher ratings for each child. Finally, the relationship between each questionnaire and the objective test measures was investigated.

The Total score from the Parent Questionnaire was correlated with age ( $r = .325$ ,  $p < .01$ ) while the Total score from the Teacher Questionnaire was not.

**Table 4** Age at which perfect performance on each task is achieved by at least 50% of children

By 6 years of age	By 7 years of age	By 8 years of age	Not by 12 years of age
First- and second-order false belief	Pictorial interpretation Referential interpretation Personalized emotions	Intentional deception Lexical interpretation Display rules to protect others' feelings Personalized thoughts	Sarcasm Display rules to protect ones' own feelings

Comparing the two versions of the questionnaire, parent's and teacher's ratings of the participants were correlated ( $r = .248, p = .019$ ).

In terms of validity, it was important to demonstrate that the participants' performance on the objective measures was related to parent and teacher reports of behaviours and abilities demonstrated in everyday living. The correlations between the parent/teacher ratings and the individual measures are presented in Table 3.

## DISCUSSION

Interaction in a social world requires more than good conversational skills and social graces. In psychology, researchers have concentrated their attention on the contribution of skills such as cognitive perspective taking, recognizing and expressing emotions, and the understanding of subtle uses of language to chart developmental changes in social interaction abilities. For each of these skills, pockets of research have documented developmental progressions within populations of typically developing children and, in some cases, yielded information on development within clinical populations (e.g., learning disabilities, autism, attention-deficit hyperactivity disorder). These latter investigations have pointed toward impairments that may exist in these areas across these different groups. Despite this progress, our ability to evaluate social interaction abilities in clinical settings remains limited. From a clinical perspective, the empirical studies are scattered: Most use experimental, non-normative tasks that are unique to each study and subject population, and few can be easily adapted for single subject assessment or clinical interpretation.

This lack of applicability is particularly troubling given that experience from clinical practice suggests that problems in social cognition (both subtle and severe) may exist in a wide range of children presenting for assessment and treatment. The limited availability of a broad-based measure of the range of skills required for engaging in successful social interaction also hampers the clinician's ability to understand a particular child's skills in relation to other areas of strength and weakness in that child's profile.

### The Tasks

The primary goal of this study was to investigate the suitability of various experimental tasks for clinical use. In accordance with the criteria we established for this study, the measures both facilitate clinical interpretation and have a relatively short administration time. Scoring details were created and/or modified to allow for greater variability in scoring and comparison of performance continuously across middle childhood. The total administration time for the protocol ranged from 30 to 50 minutes. This included all of the tasks and the screening measure (WISC-III Vocabulary subtest). Typically, younger children took longer to complete the tasks, although some older or more talkative children also required longer testing sessions.

In order to be able to evaluate single case performance, normative data for five tasks was collected for children between 6 and 12 years of age (see Appendix C). Ceiling effects across this age range were seen, as one might expect, in the false belief tasks. Children typically achieved perfect performance on these stories by 6 to 9 years

of age. Rather than excluding this portion of the task from future testing, however, it was important to maintain it in the battery for several reasons. First, given the relatively poor performance of the 6 year olds on the other tasks in this battery, it was important to establish and demonstrate that even the younger school-aged children *were* able to complete these fundamental tasks. For those in early middle childhood who may have difficulty understanding concepts such as intentional deception and sarcasm, it is important to rule out a more basic misconception about the nature of beliefs and intentions per se. The second reason for keeping this portion of the task concerns clinical utility. When testing children with cognitive impairments, poorer performance on these tasks may be noted, and it will be important to provide data for normative comparison in these cases to establish potential areas of weakness. Taken together with the intentional deception and sarcasm stories, children's performance on this overall task (including false belief stories) demonstrated good variability with steadily improving levels of performance between 6 and 12 years of age.

Ceiling effects were also apparent on several other tasks in this study despite previous research suggesting more variable performance in children between 6 and 12 years of age. On the interpretive theory of mind tasks, perfect performance was achieved by most children by 8 years of age. Still, understanding the interpretive nature of thoughts is considered an integral aspect of the social-cognitive skills children acquire during middle childhood. It is likely that this level of understanding represents a building block for more sophisticated cognitive and emotional understanding to come. For example, older children come to understand that personality traits can influence a person's point of view, or that a particular upbringing can cause an individual to react in a unique or unusual way. Understanding that different people can *in principle* have different interpretations of the same situation or stimuli is essential for these later developmental achievements. Thus, despite the ceiling effect for 8-year-olds, the task may be of value considering the impact it may have on the development of more complex skills and the importance of measuring these progressions in clinical populations.

Overall, children's performance on the stories involving display rules tended to flatten out by about 9 years of age. Unlike other tasks, this "ceiling" did not result from children achieving perfect performance at any age. At best, performance averaged near 60% on this task. As noted in the Results section, children achieved slightly higher levels of success in generating verbal display rules (70 to 80%) than facial display rules (45 to 50%) although a similar pattern of flattening around 9 to 10 years of age was still seen for both types of rules. Further investigation with a sample of older participants (e.g., adolescents) would be helpful in determining whether this plateau occurred as a function of task design or a natural stepwise pattern of development.

Finally, children's performance on tasks involving personalized thought and personalized emotion tended to reach ceiling levels around 9 years of age. Near-perfect performance on each of these tasks indicated that, by this age, children were actively considering personal information when making judgments about an individual's feelings or beliefs. The thematic content of the stories used in the Personalized Emotion task and the scenario of the Personalized Thought task were both relatively simple in order to ensure familiarity for all of the participants. It is possible that scenarios involving more sophisticated personal information (e.g., subtle personality

traits) or more mature situations may yield greater differentiation among the older participants, and are worthy of further investigation.

While the unusual ceiling effect on the Display Rules subtest may raise questions about task integrity, older children did achieve 100% performance on the majority of the other tasks, perhaps suggesting that these tasks are measuring skills that are fully developed by the middle school-age years. Still, it would be presumptuous to assume that 100% performance on these tasks suggests adult levels of social cognitive development in children as young as 9, since more complex aspects of social cognitive functioning may not have been measured by these tasks. Indeed, more sophisticated tasks have been developed for the purposes of evaluating social cognition in adults or higher functioning patient samples (e.g., individuals with Asperger Syndrome) but were not included in this study. From a clinical standpoint, these results suggest that more than one set of measures may be needed to capture the range of theory of mind development occurring between 6 and 12 years of age.

### **The Questionnaires**

Objective test performance was compared with parent and teacher ratings of children's social skills and performance in social transactions in order to establish an initial sense of ecological validity for the tasks, as well to provide an additional source of information regarding social functioning that could be used in a neuropsychological evaluation. In general, the correlations between these ratings and the clinical measures suggest a modest relationship between the objective tasks and everyday behavior.

Interestingly, these results also bear some resemblance to the executive functioning literature. It remains to be seen whether complex social behaviors which are truly defined in terms of their online performance, can be adequately captured and evaluated through laboratory tasks alone. More likely, as is the case with executive functioning, a multi-faceted evaluation of social cognitive skills will need to include elements of test completion, direct observation and observer ratings. In addition to establishing validity for these measures, the inclusion of the social-cognitive questionnaires makes available a second route of information gathering. Future studies will be needed to determine the efficacy of the questionnaire in capturing aspects of social cognitive limitations in children with reported difficulties. The questionnaire may be helpful in addressing situations where a child is able to use knowledge to pass the objective measures, but has difficulties in applying this knowledge as reported by parents and/or teachers.

### **Theory of Mind and Brain Maturation**

The relationship between normal childhood brain maturation and development of theory of mind has received limited attention (see Stuss & Anderson, 2004, for a review). Stuss and Anderson (2004) have argued that the strong correlation between executive functions and theory of mind may provide a beginning for understanding the relationship between frontal brain maturation and skill development. In the executive functioning literature, studies have documented ongoing development of executive skills from an early age until adulthood. Some have

suggested that the developmental trajectory of these skills may not be linear, but rather appears to occur in “growth spurts,” first between birth and 5 years of age, again between 7 and 9 years of age, and again between 11 and 13 years of age (see Anderson, 2002). It is postulated that these periods of skill development likely correspond to neurophysiological changes in synaptogenesis and myelination of the prefrontal cortex (see Anderson, 2002; Stuss & Anderson, 2004). With this in mind, one can begin to ask questions about the development of social cognitive skills in terms of linearity versus “growth spurts.” The spurt in theory of mind development that occurs around 4–5 years of age with the understanding of deception is well-established. The results of this study may provide very preliminary evidence for further skill development in the middle school-aged years, particularly between the ages of 7 and 10 years. It may be the case that the nature of the skills being evaluated by the tasks used in this study were not sufficient to demonstrate continued development of social cognitive skills that one might expect after 11 years of age.

## CONCLUSIONS AND FUTURE DIRECTIONS

In summary, this study yielded normative data for five tasks related to social competence that may be appropriate for administration with school-aged children. The sensitivity of these measures to deficits in clinical populations, and similarly, the ability of these tasks to differentiate among typical and atypical groups remains to be evaluated. Based on the ceiling effects around 9 to 10 years of age, further investigation of tasks and skills that develop into early adolescence seems warranted, as the scope of later social cognitive development may not be adequately captured by the measures included in this study.

Support for this project by teachers, clinicians and researchers alike, suggests that the importance of broadening our understanding of social competence is far-reaching. Unlike children with learning disabilities who are challenged academically, children with social impairments bear the burden of their disability in almost every environment. It is undeniable that one’s attitudes and opinions of another individual stem largely from social interchanges. Being at a social disadvantage puts children at risk for poor relationships with peers, parents and teachers. From an educational standpoint, increased awareness of social impairments can serve to encourage acceptance and understanding by teachers towards these challenging students.

The implications of developing this type of measure are numerous, not the least of which is the fact that interpretation guides our intervention. Depending on how one interprets inappropriate social behaviour, the treatment recommendations can be dramatically different. Enabling psychologists to measure concepts of social cognition in a more objective manner may lend weight to arguments that not all social behavior is intentional, and that specific teaching in social interaction skills is a more appropriate intervention than discipline (i.e., reward and punishment). The relationship between social interaction behaviours and brain functioning is a critical piece of knowledge that needs to be shared with families and professionals in the community.

Limitations of this study included the small sample size used to evaluate inter-rater and test-retest reliability, and the ranges of ages included, which precluded a fuller understanding of the developmental trajectories of some measures (i.e., Display Rules).

With respect to the measures specifically, future directions might include:

1. Further study of the Display Rules task in an independent sample of typically developing children in order to verify the unusual “ceiling” effect with slightly older participants and an adult control group.
2. Controlled studies of clinical populations for whom theory of mind and social cognitive deficits have been well established (i.e., autistic spectrum disorders) to determine the sensitivity of these measures in detecting known deficits.
3. A larger scale study to increase the normative data, and replicate the inter-rater reliability and test–retest reliability in large samples before the measures’ clinical utility can be established.
4. Continued investigation of issues related to ecological validity in social cognitive measures including children’s knowledge versus application of theory of mind skills. This may include further refinement and analysis of the Social Cognitive Questionnaire to determine the presence of different content areas that may reflect different elements of social cognitive functioning—much like different elements of executive functioning.

## REFERENCES

- Adams, K. & Markham, R. (1991). Recognition of affective expression by children and adolescents with and without mental retardation. *American Journal on Mental Retardation*, *96*, 21–28.
- Anderson, P. (2002). Assessment and development of executive function (EF) during childhood. *Child Neuropsychology*, *8*(2), 71–82.
- Astington, J. W., Harris, P. L., & Olson, D. R. (Eds.). (1988) *Developing theories of mind*. New York: Cambridge University Press.
- Baron-Cohen, S., Ring, H., Moriarty, J., Schmitz, B., Costa, D., & Ell, P. (1994). Recognition of mental state terms: Clinical findings in children with autism and a functional neuroimaging study of normal adults. *British Journal of Psychology*, *165*, 640–649.
- Baron-Cohen, S., Tager-Flusberg, H., & Cohen, D. J. (2000.) *Understanding other minds: Perspectives from developmental cognitive neuroscience* (2nd ed.). Great Britain: Oxford University Press.
- Brothers, L. (1990). The social brain: A project for integrating primate behaviour and neurophysiology in a new domain. *Concepts in Neuroscience*, *1*, 27–51.
- Carpendale, J. I. & Chandler, M. J. (1996). On the distinction between false-belief understanding and subscribing to an interpretive theory of mind. *Child Development*, *67*, 1686–1706.
- Chandler, M. J. (1988). Doubt and developing theories of mind. In J. W. Astington, P. L. Harris, & D. R. Olson (Eds.), *Developing theories of mind* (pp. 387–413). New York: Cambridge University Press.
- Chandler, M. J. & Helm, D. (1984). Developmental changes in the contribution of shared experience to social role taking competence. *International Journal of Behavioural Development*, *7*, 145–156.
- Chandler, M. J. & Lalonde, C. E. (1996). Shifting to an interpretive theory of mind: 5- to 7-year-olds’ changing conceptions of mental life. In A. J. Sameroff & M. M. Haith (Eds.), *The five to seven year shift: The age of reason and responsibility* (pp. 111–139). Chicago, IL: The University of Chicago Press.
- Demorest, A., Meyer, C., Phelps, E., Gardner, H., & Winner, E. (1984). Words speak louder than actions: Understanding deliberately false remarks. *Child Development*, *55*, 1527–1534.

- Dennis, M., Barnes, M. A., Wilkinson, M., & Humphreys, R. P. (1998). How children with head injury represent real and deceptive emotion in short narratives. *Brain and Language, 61*, 450–483.
- Gallagher, H. L. & Frith, C. D. (2003). Functional imaging of “theory of mind.” *Trends in Cognitive Neurosciences, 7*(2), 77–83.
- Gnepp, J. & Gould, M. E. (1985). The development of personalized inferences: Understanding other people’s emotional reactions in light of their prior experiences. *Child Development, 56*, 1455–1464.
- Gnepp, J. & Hess, D. L. R. (1986). Children’s understanding of verbal and facial display rules. *Developmental Psychology, 22*, 103–108.
- Greene, R. W., Biederman, J., Faraone, S. V., Ouellette, C. A., Penn, C., & Griffin, S. M. (1996). Toward a new psychometric definition of social disability in children with attention-deficit hyperactivity disorder. *Journal of the American Academy of Child and Adolescent Psychiatry, 35*, 571–578.
- Gregory, C., Lough, S., Stone, V., Erzinclioğlu, S., Martin, L., Baron-Cohen, S., et al. (2002). Theory of mind in patients with frontal variant frontotemporal dementia and Alzheimer’s disease: Theoretical and practical implications. *Brain, 125*(Pt 4), 752–764.
- Grezes, J., Frith, C. D., & Passingham, R. E. (2004). Inferring false beliefs from the actions of oneself and others: An fMRI study. *Neuroimage, 21*(2), 744–750.
- Hala, S., Chandler, M. J., & Fritz, A. S. (1991). Fledgling theories of mind: Deception as a marker of three-year-olds’ understanding of false belief. *Child Development, 62*, 83–97.
- Happé, F. (1993). Communicative competence and theory of mind in autism: A test of relevance theory. *Cognition, 48*, 101–119.
- Harris, P. L., Donnelly, K., Guz, G. R., & Pitt-Watson, R. (1986). Children’s understanding of the distinction between real and apparent emotion. *Child Development, 57*(4), 895–909.
- Jackson, H. F. & Moffatt, N. J. (1987). Impaired emotional recognition following severe head injury. *Cortex, 23*, 293–300.
- Lalonde, C. E. & Chandler, M. J. (1995). False belief understanding goes to school: On the social-emotional consequences of coming early or late to a first theory of mind. *Cognition and Emotion, 9*(2–3), 167–185.
- Lalonde, C. E. & Chandler, M. J. (2002). Children’s understanding of interpretation. *New Ideas in Psychology, 20*(2–3), 163–198.
- Leekham, S. R. (1991). Jokes and lies: Children’s understanding of intentional falsehood. In A. Whiten (Ed.), *Natural theories of mind: Evolution, development and simulation of everyday mindreading* (pp. 159–174). Oxford, England: Basil Blackwell.
- Leekham, S. R. & Perner, J. (1991). Does the autistic child have a metarepresentational deficit? *Cognition, 40*, 203–218.
- Matthys, W., Cuperus, J. M., & van Engeland, H. (1999). Deficient social problem-solving in boys with ODD/CD, with ADHD, and with both disorders. *Journal of the American Academy of Child and Adolescent Psychiatry, 38*, 311–321.
- McEvoy, R. E., Rogers, S. J., & Pennington, B. F. (1993). Executive function and social communication deficits in young autistic children. *Journal of Child Psychology and Psychiatry, 34*, 563–578.
- Milch-Reich, S., Campbell, S. B., Pelham, Jr., W. E., Connelly, L. M., & Geva, D. (1999). Developmental and individual differences in children’s on-line representations of dynamic social events. *Child Development, 70*, 413–431.
- Moffatt, C. W., Hanley-Maxwell, C., & Donnellan, A. M. (1995). Discrimination of emotion, affective perspective-taking and empathy in individuals with mental retardation. *Education and Training in Mental Retardation and Developmental Disabilities, 30*, 76–85.
- Montgomery, D. E. (1993). Young children’s understanding of interpretive diversity between different-age listeners. *Developmental Psychology, 29*(2), 337–345.

- O'Neill, D. K. & Gopnik, A. (1991). Young children's understanding of the sources of their knowledge. *Developmental Psychology*, 27, 390–397.
- Ozonoff, S., Pennington, B., & Rogers, S. J. (1991). Executive function deficits in high functioning autistic children: Relationship to theory of mind. *Journal of Child Psychology and Psychiatry*, 31, 1081–1105.
- Perner, J. (1988). Higher-order beliefs and intentions in children's understanding of social interaction. In J. W. Astington, P. Harris & D. Olson (Eds.), *Developing theories of mind* (pp. 271–294). Cambridge: Cambridge University Press.
- Petterson, L. (1991). Sensitivity to emotional cues and social behaviour in children and adolescents after head injury. *Perceptual and Motor Skills*, 73, 1139–1150.
- Rogers, S. J. & Pennington, B. F. (1991). A theoretical approach to the deficits in infantile autism. *Development and Psychopathology*, 3, 137–162.
- Rourke, B. P. & Tsatsanis, K. D. (1996). Syndrome of nonverbal learning disabilities: Psycholinguistic assets and deficits. *Topics in Language Disorders*, 16, 30–44.
- Saarni, C. (1984). An observational study of children's attempts to monitor their expressive behaviour. *Child Development*, 55(4), 1504–1513.
- Snowden, J. S., Gibbons, Z. C., Blackshaw, A., Doubleday, E., Thompson, J., Craufurd, D., et al. (2003). Social cognition in frontotemporal dementia and Huntington's disease. *Neuropsychologia*, 41(6), 688–701.
- Sodian, B. (1991). The development of deception in young children. *British Journal of Developmental Psychology*, 9, 173–188.
- Sprouse, C. A., Hall, C. W., Webster, R. E., & Bolen, L. M. (1998). Social perception in students with learning disabilities and attention deficit/hyperactivity disorder. *Journal of Nonverbal Behavior*, 22, 125–134.
- Stone, V. E., Baron-Cohen, S., & Knight, R. T. (1998). Frontal lobe contributions to theory of mind. *Journal of Cognitive Neuroscience*, 10(5), 640–656.
- Stuss, D. T. & Anderson, V. (2004). The frontal lobes and theory of mind: Developmental concepts from adult focal lesion research. *Brain and Cognition*, 55, 69–83.
- Taylor, M., Cartwright, B. S., & Bowden, T. (1991). Perspective taking and theory of mind: Do children predict interpretive diversity as a function of differences in observers' knowledge? *Child Development*, 62, 1334–1351.
- Terwogt, M. M. & Olthof, T. (1989). Awareness and self-regulation of emotion in young children. In C. Saarni & P. Harris (Eds.), *Children's understanding of emotion* (pp. 209–237). Cambridge: Cambridge University Press.
- Terwogt, M. M. & Rieffe, C. (2003). Stereotyped beliefs about desirability: Implications for characterizing the child's theory of mind. *New Ideas in Psychology*, 21(1), 69–84.
- Wechsler, D. (1991). *The Wechsler intelligence scale for children* (3rd ed.). San Antonio, TX: Psychological Corporation.
- Wimmer, H. & Perner, J. (1983). Beliefs about beliefs: Representation and constraining function of wrong beliefs in young children's understanding of deception. *Cognition*, 13(1), 103–128.
- Zelazo, P. D. & Mueller, U. (2002). Executive functioning in typical and atypical children. In U. Goswami (Ed.), *Blackwell Handbook of Childhood Cognitive Development* (pp. 445–469). Oxford: Blackwell.
- Zeman, J. & Garber, J. (1996). Display rules for anger, sadness, and pain: It depends on who is watching. *Child Development*, 67, 957–973.

## APPENDIX A: TASK EXAMPLES

### 1. False Belief (Second Order)

The grade three class is going to get a new pet. A little bunny named Whiskers. The teacher, Mrs. James, wants the bunny to be a surprise. She tells the class they will be getting a turtle. She hides the bunny in the coatroom so that she can surprise the class after lunch. The bell rings, and one of the students, Sally, goes to the coatroom to get her skipping rope. She sees the bunny. Sally thinks, "We are not getting a turtle for our class pet, we are getting a bunny!" Then she goes to have lunch. Mrs. James does not see Sally find the bunny. At lunch, the principal says to Mrs. James, "Do any of your students know you got a bunny for the class?"  
Memory Questions: What new pet did Mrs. James get for her class?

Did Mrs. James see Sally find the bunny?

Prediction Question: What does Mrs. James say to the principal?

Explanation Question: Why does she say that?

### 2. Intentional Deception

Tina got a new pair of rollerblades for her birthday. Tina does not know how to skate, but she wants to learn. Everyday Tina goes outside to practice on the street in front of her house. Tina is not very good at skating. She falls down a lot. Tina's friend Jeff is outside playing with his yo-yo. Jeff sees Tina fall down many times. Jeff walks over to talk to Tina. Tina says to Jeff, "I have been practicing my skating everyday. Can you tell?" Jeff smiles and says, "Yeah, you are pretty good."

Memory Question: What happened in the story?

Question 1: Did Jeff think Tina was a good rollerblader or not?

Question 2: How do you know?

Question 3: Did Jeff want Tina to think that she was a good rollerblader or not?

Question 4: How do you know he wanted Tina to think that?

### 3. Sarcasm

Over the summer Daniel grew three centimeters. Some of his old clothes do not fit him anymore. One morning Daniel gets dressed and sees that his blue pants are too short. His socks are sticking out from the bottom of his pants. Daniel wears the blue pants to school anyway, because he will not have time to get new pants until the weekend. At school, Daniel sits next to Sarah. Sarah looks at Daniel's short pants. She giggles and points at the pants. Sarah says, "Those pants are very stylish."

Memory Question: What happened in the story?

Question 1: Did Sarah think that Daniel's pants were nice or not?

Question 2: How do you know she thought that?

Question 3: Did Sarah want Daniel to think she liked his pants or not?

Question 4: How do you know she wanted Daniel to think that?

#### 4. Display Rules

This is a story about David. Last week it was David's birthday. David's grandmother came for dinner and gave David his birthday present. She was there when David opened it. It was a hat. David thought it was very ugly.

What did he say? Why did he say that?

What kind of face did he make? Why did he make that face?

#### 5. Personalized Emotions

This is a story about Donna. One night, on television, Donna watched a movie about poisonous snakes that live in the jungle. The snakes had large mouths and giant, sharp teeth. The next day, Donna's teacher announced to the class that they were going to have a new class pet: A snake named Fang!

How does Donna feel? Why?

### APPENDIX B: SAMPLE ITEMS FROM THE PARENT/TEACHER QUESTIONNAIRES

- Does your child (student) understand the meaning of a sarcastic remark?
- Does your child make appropriate eye contact during conversations?
- Does your child try to win an argument by persuading someone to change his/her mind?
- Does your child understand the difference between something that is done "on purpose" versus "not on purpose"?
- Does your child understand that a person can be happy and sad about something at the same time?
- Does your child acknowledge that two people can have different opinions on the same issue?
- Is your child able to tell how others are feeling by "reading" their facial expressions?
- Is your child able to figure out how a person might be feeling from their body language (e.g., posture)?
- Does your child understand when s/he is being teased?
- Does your child realize that a person may feel a certain way "on the inside" but display a different emotion "on the outside," in particular, when expression of one's true feelings is not socially acceptable, may hurt others, or may cause shame or embarrassment to oneself.

### APPENDIX C: PRELIMINARY NORMATIVE DATA

Percent correct achieved on each task by age group (scores reflect % correct)

Age group (Years)	Task											
	False Belief		Interpretive TOM		Display Rules		Personalized Thoughts		Personalized Emotions			
	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>		
6	56.25	17.27	46.30	28.75	15.63	20.73	54.17	23.44	58.33	28.87		
7	66.07	10.32	79.37	18.93	21.43	25.21	57.14	22.85	76.19	27.51		
8	62.50	10.21	85.56	18.92	51.25	25.99	85.00	24.15	90.00	16.10		
9	77.94	15.01	93.46	13.64	58.82	23.70	82.35	22.99	96.08	11.07		
10	75.00	18.99	94.44	8.44	51.79	22.39	78.57	25.68	88.10	28.06		
11	82.95	14.00	95.96	8.99	59.09	16.86	88.64	20.51	96.97	10.05		
12	85.23	9.87	96.97	7.19	62.50	25.62	79.55	24.54	96.97	10.05		