



Developing an understanding of others' emotional states: Relations among affective theory of mind and empathy measures in early childhood

Caitlyn M. M. Gallant* , Lydia Lavis and Caitlin E. V. Mahy 

Department of Psychology, Brock University, St. Catharines, Ontario, Canada

Theory of mind (ToM) consists of cognitive and affective components; however, few studies have evaluated the coherence of affective ToM measures and their associations with empathy. This research examined the relations among affective ToM tasks and assessments of empathy, measured directly and via parent reports in 4- to 6-year-olds. Children ($N = 117$) completed: an Appearance-Reality Emotion Task, an adapted Reading the Mind in the Eyes Test, an Affective Stories Task, and an empathy task. Parents reported on children's ToM and empathy, and language was assessed using a Picture Vocabulary Test. Controlling for language, no relationships were found among affective ToM measures and children's age was only related to the Affective Stories Task. Further, controlling for age, only parent-reported empathy was associated with the Appearance-Reality Emotion Task. Once vocabulary and age were controlled, measures of affective ToM are unrelated and different developmental patterns emerged across measures.

Statement of contribution

What is already known on this subject

- Affective theory of mind is a distinct subcomponent of theory of mind that corresponds to an independent developmental mechanism.
- However, little research has been conducted on affective ToM, its developmental trajectory during the preschool years, its assessment, and its relationship with related constructs, such as empathy.

What the present study adds

- Children's performance on affective ToM tasks was unrelated once age and language abilities were accounted for. Thus, there is a need to examine affective ToM and its measurement more extensively to ensure we are effectively capturing this construct.
- This study was the first to establish a Preschool Reading the Mind in the Eyes Task that uses images of children and pictorial response options and an Affective Stories Task that captures age-related changes in affective ToM beyond language skills.

The ability to infer others' mental states, such as their thoughts, desires, beliefs, and emotions, is known as theory of mind (ToM; Baron-Cohen, Tager-Flusberg, & Cohen, 1994). ToM is a fundamental component of social cognition, as it enhances our ability to anticipate others' behaviours and promotes cooperation, selflessness, and interpersonal

*Correspondence should be addressed to Caitlyn M. M. Gallant, Department of Psychology, Brock University, 1812 Sir Isaac Brock Way, St. Catharines, ON L2S 3A1, Canada (email: cg14xx@brocku.ca).

sensitivity (Slaughter, Dennis, & Pritchard, 2002). Disorders characterized by deficits in ToM, such as autism spectrum disorder (ASD), underscore the importance of this ability, as individuals with ASD often misinterpret social interactions and are subsequently perceived as insensitive or inappropriate (Castelli, Frith, Happé, & Frith, 2002; Klin, Jones, Schultz, Volkmar, & Cohen, 2002). Among children, ToM is equally important—it has been associated with greater popularity, academic achievement, and more sophisticated interpersonal strategies (e.g., Blair & Razza, 2007; Caputi, Lecce, Pagnin, & Banerjee, 2012). Thus, ToM plays a critical role in children's interpersonal interactions and everyday social behaviour.

According to Wellman and Liu (2004), children's ToM follows a reliable trajectory; children first develop an understanding of diverse desires, followed by diverse beliefs, knowledge access, false beliefs, and hidden emotions. Advancement through each of these stages is thought to occur as children collect evidence about mind-behaviour relations, develop a theory, and revise/update their theory based on incompatible evidence (i.e., theory-theory account of ToM; Gopnik, 2003). Accordingly, this perspective assumes that earlier stages of reasoning serve as prerequisites for later stages and that each development reflects an advancement in ToM. However, there is growing support for the idea that ToM is not unidimensional and that these developments may capture different types of ToM reasoning. According to Warnell and Redcay (2019), several ToM measures are uncorrelated when controlling for potential confounds (e.g., age, verbal abilities), implying that they do not share a common underlying factor. Moreover, recent investigations show that ToM consists of independent cognitive and affective components that correspond to independent mechanisms and neural substrates (e.g., Kalbe *et al.*, 2010; Sebastian *et al.*, 2011). Therefore, there is a need for more targeted investigations examining different aspects of ToM and their measurement, rather than ToM as an overarching construct. The current study contributes to this underexplored topic by assessing affective ToM measures and their associations in preschool-aged children.

Affective ToM involves understanding others' emotional states and is thought to rely on the ventromedial prefrontal cortex (e.g., Shamay-Tsoory, Tomer, Berger, Aharon-Peretz, & Goldsher, 2005). In particular, it has been proposed that children understand others' affective states by engaging in simulation (Gallese & Goldman, 1998; Harris, 1992)—that is, children rely on their own physiological/emotional states to mimic others' emotions in the self (Kalbe *et al.*, 2007). Accordingly, it has been demonstrated that affective ToM is closely associated with children's empathic abilities (e.g., Ginsburg *et al.*, 2003; Shamay-Tsoory, Tomer, Goldsher, Berger, & Aharon-Peretz, 2004). For example, in 6-year-olds, Strayer and Roberts (1989) found that the ability to label a target character's emotion was positively associated with self- and parent-reported empathic abilities. Similarly, performance on the Reading the Mind in the Eyes Test (RMET; Baron-Cohen *et al.*, 2001), an index of affective ToM, has been shown to positively correlate with measures of empathy (e.g., Carroll & Chiew, 2006; Lawrence, Shaw, Baker, Baron-Cohen, & David, 2004), although many have asserted that the RMET measures emotion recognition, rather than affective ToM (e.g., Oakley, Brewer, Bird, & Catmur, 2016).

Nonetheless, few ToM studies have focused on the development and assessment of affective ToM *during preschool*. To date, a range of behavioural tasks and parent reports have been employed; however, the coherence of these measures and the degree to which they reflect a common underlying construct remain unclear. Indeed, as Warnell and Redcay (2019) have highlighted, the vast majority of ToM studies employ only a single measure of ToM and thereby suggest that all ToM measures capture the same construct. In

particular, children's affective ToM has predominantly been assessed using hidden-emotion paradigms (e.g., Pons & Harris, 2000; Wellman & Liu, 2004) and complex affect recognition tests originally designed for adult and/or clinical populations (e.g., RMET; Baron-Cohen *et al.*, 2001). In these hidden-emotion tasks, children are asked to evaluate how a protagonist feels in various situations. Often, these tasks require an understanding that individuals can feel one emotion on the inside; yet, they can display different emotions to conceal their true feelings. Thus, their interpretation is often obscured by the additional cognitive and verbal demands they create (see Eisenberg, Spinrad & Sadovsky, 2006). For instance, children are often required to maintain and manipulate information in working memory or asked to infer complex emotions that require superior verbal abilities (e.g., jealous, guilty), making it difficult to decipher whether task impairments reflect immature affective ToM or other cognitive limitations. Further, it has been argued that an understanding of display rules—the set of rules that define the appropriate expression of emotion in particular contexts (Ekman & Friesen, 1975)—might limit children's performance on hidden-emotion tasks despite developed affective perspective taking. Taken together, there is a need to evaluate the coherence and validity of existing affective ToM measures and how they correlate with related processes (e.g., Völlm *et al.*, 2006). At present, this gap in the literature makes it difficult to establish the developmental trajectory of affective ToM among preschoolers.

In the emotion understanding literature, 3-year-olds can discriminate between happiness, anger, and sadness and show an understanding of situational antecedents of emotion (e.g., Cutting & Dunn, 1999; Harris, 1989). Similar to Wellman and Liu's ToM (2004) scaling, Pons, Harris, and de Rosnay (2004) established that children's emotion understanding develops through three developmental periods between the ages of three and 11 years: (1) Children develop an understanding of the external causes/cues of emotions and their visible expression from 3 to 5 years, (2) children comprehend the mentalistic aspects of emotions and their associations with internal states between 5 and 7 years, and (3) children understand mixed emotions, the link between morality and emotions, and how emotions may be regulated by cognition between 7 and 11 years. This research has important implications for the development of affective ToM, as it suggests that preschoolers can reason about others' feelings well before they master hidden emotions and other types of ToM. In fact, as noted by Wellman and Liu (2004), understanding of emotions is found to precede false-belief understanding when it is measured by multifaceted tasks such as Denham's test (1986). Despite these findings, many ToM researchers have suggested that affective ToM continues to develop well beyond the preschool years (Shamay-Tsoory *et al.*, 2005; Wellman & Liu, 2004). For instance, in an investigation by Sebastien *et al.* (2011), adolescents were found to make significantly more affective ToM errors compared to adults, indicating that emotion understanding may have an extended developmental trajectory.

From a theoretical perspective, therefore, there is a need to integrate the ToM and emotion understanding literatures in the context of affective ToM development; thus, the primary purpose of this research was to examine the development and assessment of affective ToM in 4- to 6-year-old children. In addition, we sought to investigate the associations among different measures of affective ToM and empathy as an index of convergent validity. Given that the RMET relies heavily on children's verbal comprehension (e.g., Ahmed & Miller, 2011; Stanford, Messinger, Malaspina, & Corcoran, 2011), an adapted version was created (the Preschool RMET). Notably, this adapted version is one of the first attempts to use images of children, rather than adults, and pictorial response options. Further, since a variety of tasks has been used to capture affective ToM reasoning,

we employed a diverse set of measures that differed in modality, complexity (i.e., verbal or cognitive requirements), and the method of measurement (i.e., children's behavioural performance and parent report). Moreover, to ensure that age-related patterns reflect true advancements in affective ToM and not increasing language abilities, we also statistically controlled for the effects of language ability. It was hypothesized that: (1) Children's performance on all measures of affective ToM and parent reports of ToM would be intercorrelated, even when holding verbal abilities constant; (2) age would predict superior affective ToM across all tasks and parent reports after controlling for verbal abilities; and (3) affective ToM measures would be positively associated with parent ratings and child measures of empathy, when controlling for age.

Method

Participants

Typically developing children ($N = 123$) between the ages of 4 and 6 years participated in this study. Six participants were excluded for being younger than four ($n = 2$), not participating in any of the tasks ($n = 3$), and for parent interference in procedure (i.e., prompting child's responses; $n = 1$). The final sample consisted of 117 children: 41 four-year-olds (21 girls; $M_{\text{age}} = 51.83$ months, $SD = 4.04$), 41 five-year-olds (24 girls; $M_{\text{age}} = 65.37$ months, $SD = 3.70$), and 35 six-year-olds (17 girls; $M_{\text{age}} = 77.23$ months, $SD = 3.34$). Participants were recruited from the Ontario Science Centre in Toronto, Canada. Children came from diverse ethnic backgrounds (37% Asian, 32% Caucasian/white, 14% mixed background, 3% Hispanic, 3% Asian Indian, 3% Middle Eastern, 2% Black, and 6% others) and were mainly middle class (80% of family incomes above 40,000 dollars per year).

Materials

Child measures

Affective Stories Task (AST). Children were read aloud five affective stories involving a character named Jamie. Each story was accompanied by a cartoon image where Jamie's face was left blank. Children were asked to select the emotion (from five cartoon faces; e.g., happy, sad, scared, angry, or surprised) that best depicts how Jamie feels in each story. All stories were 1 to 2 sentences in length (based on those compiled by Garner, Jones, & Miner, 1994; see Table 1) and each child experienced the stories in the same order, resulting in a total score out of five. Similar to Component II of the Test of Emotion Comprehension (Pons *et al.*, 2004), the AST required children to make an emotional state attribution based on external/situational causes by pointing to a response option. Data from this task were excluded for three children because they did not understand the task ($n = 2$) or failed to complete the task ($n = 1$).

Preschool Reading the Mind in the Eyes Test (RMET). In the Preschool RMET (adapted from Baron-Cohen *et al.*, 2001), children were asked to infer how another individual was feeling based on photographs of the eye region of the face. Specifically, the task consisted of eight black-and-white photographs (derived from the NIMH Child Emotional Faces Picture Set; Egger *et al.*, 2011) of children depicting different emotions, accompanied by four emotional cartoon response options (used in place of words to reduce verbal

Table 1. List of stories and their associated emotions in the affective stories task

Emotion	Story
Happy	It is Jamie's birthday. Jamie is given a party with lots of cake and fun games to play and presents too
Sad	Jamie couldn't play a game very well and some of the other kids laughed at Jamie
Scared	Jamie was dreaming about a monster. Jamie had a nightmare
Surprised	Jamie came home from school one day. Jamie's mom had pink hair
Mad	Jamie was building a big tower of blocks. Jamie's friend came along and knocked it over

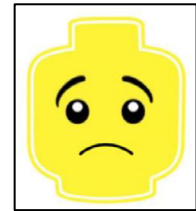
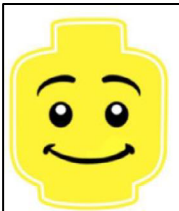
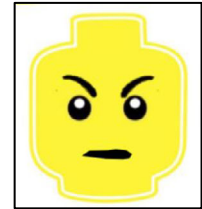
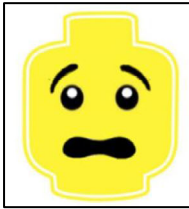


Figure 1. Sample item of the Preschool Reading the Mind in the Eyes Task (adapted from Baron-Cohen, Wheelwright, Spong, Scahill, & Lawson, 2001; created using photographs from the NIMH Child Emotional Faces Picture Set; Egger et al., 2011).

demands; see Figure 1). The photographs consisted of an equal number of male and female child actors, ranging from 10- to 17-year-olds. The four response options (happy, sad, mad, and scared) were consistent across items, but their location around the four corners of the photograph differed for each trial to avoid any potential location biases. Children were asked to select the emotion that best described how the photographed individual was feeling. At the start of each trial, the experimenter verbally listed each emotion while pointing to the response options. Children were permitted to provide both verbal and pointing responses for each item; however, the majority provided verbal responses. Consistent with the instructions of the child RMET (Baron-Cohen *et al.*, 2001), participants were instructed to guess if they were unsure. A practice trial was

administered, and the majority of children correctly answered this trial. Children's RMET total score ranged from 0 to 8. Two children were excluded because they failed to complete the task.

Empathy task. In this task (adapted from Zahn-Waxler, Robinson, & Emde, 1992), the experimenter pretended to clip her finger in a clipboard to examine children's empathic reactions. The experimenter clipped the clipboard loudly and simulated pain for a duration of 20 s by verbally exclaiming that she hurt herself and by holding her finger. The task concluded with the experimenter indicating that she felt better and was no longer hurt. Empathic concern was scored across three categories (affective responses, verbal empathic initiations, and behaviour activation) using a 4-point scale (adapted from Butean, Costescu, & Dobrean, 2014), which resulted in a maximum score of 12. Inter-rater reliability on 30% of the data was substantial ($\kappa = .610$). Seventeen children were excluded from analyses due to: equipment failure ($n = 13$), unsuccessful pain simulation (i.e., the child was too suspicious; $n = 2$), and failure to complete the task ($n = 2$).

Appearance-Reality Emotion Task (A-RE). In the A-RE Task (Wellman & Liu, 2004), children were told a story about a character who receives an unwanted gift instead of the gift he was promised. The children were told that the character must hide his true feelings to receive gifts in the future, and they were asked to distinguish between the character's true feelings and those visibly expressed (e.g., 'happy, sad, or okay?'). Children were provided with a visual scale depicting a happy face, a sad face, and a neutral face as response options and were trained to use the scale. Responses were considered correct if they indicated that the character's experienced emotion was more negative than the character's expressed emotion (scored as pass or fail). Prior to the two test questions, children were asked two comprehension control questions—those who answered these questions incorrectly were read the story again (up to a maximum of three times) and were then administered the test questions regardless of whether they passed the control questions on the third trial. Overall, 113 children completed the A-RE Task: 28 children failed both control questions (25%), 35 children passed only one of the control questions (31%), while 50 children passed both control questions (43%). Only those who answered both control questions were included in analyses.

Picture Vocabulary Test (PVT). To control for language abilities, children were administered the PVT from the NIH Toolbox on an iPad as an index of receptive vocabulary (Zelazo & Bauer, 2013). The PVT requires children to point to one of four pictures that match a word (e.g., 'vehicle' and 'squash'). Twelve children were excluded from analyses because they failed to complete the task. Raw uncorrected PVT scores (range = 34–81) were used for all analyses.

Parent measures

The Empathy Questionnaire (EmQue). The EmQue (Rieffe, Ketelaar, & Wiefferink, 2010) is a parent report form that consists of 20 items representing three facets of empathy: (1) Emotion Contagion, (2) Attention to Others' Feelings, and (3) Prosocial

Table 2. Final list of items comprising the Emotion Subscale from the Children's Social Understanding Scale (CSUS)

Number	Item
2	Tries to understand the emotions of other people (e.g., wants to know why you are crying)
8	Talks about how people feel (e.g., 'I'm happy'; 'She's angry')
20	Realizes that if s/he does something bad, others may get mad
24	When given an undesirable gift, pretends to like it so as not to hurt the other person's feelings
27	Understands that different people can have different feelings about the same thing (e.g., one child likes a dog but another child is scared of it)
33	Talks about conflicting emotions (e.g., 'I am happy to go on vacation, but I am sad about leaving friends behind')

Actions. Parents were asked to rate the frequency of each statement over the past 2 months on a 3-point scale (e.g., 0 = never, 1 = sometimes, 2 = often). Mean scores based on all items were used for analyses, resulting in a maximum possible score of 2. Internal consistency for the EmQue was acceptable (Cronbach's $\alpha = .77$; Peterson, 1994).

Children's Social Understanding Scale (CSUS). The CSUS (Tahiroglu *et al.*, 2014) is a 42-item parent report of ToM which consists of six subscales that capture different aspects of ToM. The Emotion subscale score (1-4) was used as a parent rating of affective ToM. Inspection of interitem correlations revealed that one Emotion subscale item (item 41) was weakly correlated with the other items and so it was deleted. Thus, a 6-item Emotion subscale was created (Cronbach's $\alpha = .62$; see Table 2).

Procedure

Parents gave written informed consent, and children gave verbal assent. Parents then completed a brief demographic questionnaire, the CSUS, and the EmQue, while their children participated in the study in a quiet room at the Science Centre. Children were tested individually for 20 min, and tasks were administered in the following fixed order: the AST, the Preschool RMET, the Empathy Task, the A-RE Task, and the PVT. At the end of the session, families were thanked for their participation, and children received a sticker and a certificate. All procedures were approved by the Brock University Research Ethics Board.

Results

Preliminary analyses

Preliminary analyses demonstrated that sex had no significant effect on children's performance on the AST, Preschool RMET, and A-RE Task, nor did it predict scores on the Empathy task and parent reports on the CSUS and EmQue ($ps > .06$). Bivariate correlations between affective ToM measures and children's verbal abilities were significant ($rs > .25$, $ps < .01$) except for the relation between verbal ability and the Preschool RMET, $r(106) = .15$, $p = .13$. Thus, PVT scores were entered as a covariate in all analyses to examine age-related changes in affective ToM independent of verbal abilities.

Relations among affective ToM measures

Children's performance on the AST significantly correlated with scores on the Preschool RMET, $r(112) = .20$, $p = .04$, and parent ratings on the Emotion subscale of the CSUS, $r(113) = .25$, $p = .01$, before controlling for verbal abilities. Further, scores on the Preschool RMET and parent ratings on the Emotion subscale of the CSUS were positively correlated, $r(114) = .19$, $p = .038$; however, these relations were no longer significant after accounting for PVT scores. No other relations were significant (see Table 3).

Affective ToM and age

Given that the correlational analyses revealed weak relations among the affective ToM measures, age-related changes in affective ToM reasoning were examined for each measure independently (see Table 4). Multiple hierarchical regression results indicated

Table 3. Correlations among age, affective theory of mind tasks, and parent reports of theory of mind

Variables	1	2	3	4	5
1. Age	–				
2. AST	.39** (.21*)	–			
3. P-RMET	.19* (.20)	.20* (.15)	–		
4. A-RE	.29* (.12)	.15 (.01)	.17 (.13)	–	
5. CSUS Emotion	.27** (.14)	.25** (.15)	.19* (.16)	–.07 (–.18)	–

Note. A-RE = Appearance-Reality Emotion Task (only children who passed both control questions; values represent point biserial correlations); AST = Affective Stories Task; CSUS = Children's Social Understanding Scale; P-RMET = Preschool Reading the Mind in the Eyes Task.

Partial correlations, controlling for language abilities (performance on the Picture Vocabulary Test [PVT]), are shown in parentheses. *ns* ranged from 42 to 117. Age = child's age in months (calculated using the child's date of birth and the date of testing).

* $p < .05$; ** $p < .01$.

Table 4. Multiple hierarchical regressions with affective theory of mind tasks regressed on verbal abilities and age

	Model 1			Model 2		
	<i>B</i>	<i>SE B</i>	β	<i>B</i>	<i>SE B</i>	β
AST						
PVT	.05	.01	.42***	.03	.01	.26*
Age				.03	.01	.30**
P-RMET						
PVT	.02	.01	.15	.01	.02	.05
Age				.02	.02	.18
CSUS Emotion						
PVT	.01	.00	.29**	.01	.01	.21
Age				.01	.01	.14

Note. Age = child's age in months (calculated using the child's date of birth and the date of testing); AST = Affective Stories Task; CSUS = Children's Social Understanding Scale; P-RMET = Preschool Reading the Mind in the Eyes Task; PVT = Picture Vocabulary Test. *ns* ranged from 42 to 117.

* $p < .05$; ** $p < .01$; *** $p < .001$.

Table 5. Descriptive statistics for affective ToM and empathy measures across age groups

Measure	4-year-olds	5-year-olds	6-year-olds	α	N
	(<i>n</i> = 41) <i>M</i> (<i>SD</i>)	(<i>n</i> = 41) <i>M</i> (<i>SD</i>)	(<i>n</i> = 35) <i>M</i> (<i>SD</i>)		
AST	2.34 (1.19)	3.32 (1.01)	3.54 (1.09)	–	114
P-RMET	5.28 (1.74)	6.07 (1.19)	5.85 (1.40)	–	115
A-RE	4 (10%)	13 (32%)	22 (63%)	–	50
CSUS Emotion	3.14 (0.50)	3.29 (0.46)	3.46 (0.36)	.62	117
Empathy task	5.00 (1.51)	5.03 (1.40)	5.29 (1.74)	–	100
EmQue	0.99 (0.26)	1.12 (0.28)	1.04 (0.24)	.77	117
PVT	56.08 (9.81)	63.81 (6.87)	70.37 (8.88)	–	105

Note. A-RE = Appearance-Reality Emotion Task (*n* (%) of children in each age group who were correct; only children who passed both control questions); AST = Affective Stories Task (scores ranged from 0 to 5); CSUS = Children's Social Understanding Scale; Empathy Task (scores ranged from 3 to 10); EmQue = Empathy Questionnaire; *M* = mean; P-RMET = Preschool Reading the Mind in the Eyes Task (scores ranged from 0 to 8); PVT = Picture Vocabulary Test; *SD* = standard deviation.

that age in months significantly predicted accuracy on the AST, over and above the effects of verbal abilities, $b = .03$, $p < .01$, 95% CI [0.01, 0.05], whereby older children made more accurate emotion attributions relative to younger children. Moreover, after controlling for verbal abilities, age did not significantly predict Preschool RMET performance, $b = .02$, $p = .13$, 95% CI [–0.01, 0.05], nor was it a significant predictor of parent reports of affective ToM on the CSUS, $b = .01$, $p = .23$, 95% CI [–0.00, 0.02]. When examining only those who passed both control questions on the A-RE Task ($n = 50$), the logistic regression model with all predictors was statistically significant, $\chi^2(2) = 7.58$, $p < .05$; however, neither verbal abilities nor age in months independently predicted the odds of being correct on the A-RE Task, $p > .05$. Table 5 shows means and standard deviations for all measures by age group and the number of children in each age group who passed the A-RE Task.

Relations among affective ToM and empathy measures

Table 6 shows that the A-RE Task was negatively and significantly correlated with scores on the EmQue, while the Emotion subscale of the CSUS was positively and significantly related to EmQue scores.¹ Notably, only the associations with the EmQue persisted when holding age constant—all relations among the affective ToM measures were also no longer significant.

Discussion

Affective ToM measurement

Consistent with Warnell and Redcay (2019), the current study demonstrated that many of the relations among affective ToM measures disappear once controlling for language or

¹ When the EmQue subscales were examined separately, parent ratings on the CSUS Emotion subscale were found to be positively correlated with the Attention to Others' Feelings subscale and with the Prosocial Actions subscale (but not the Emotion Contagion subscale). However, none of the EmQue subscales were independently associated with the A-RE Task. Similar results were found when controlling for PVT scores.

Table 6. Correlations among affective theory of mind and empathy measures controlling for age

Variables	1	2	3	4	5	6
1. AST	–					
2. P-RMET	.20* (.14)	–				
3. A-RE	.15 (.04)	.17 (.13)	–			
4. CSUS Emotion	.25** (.15)	.19* (.16)	–.07 (–.18)	–		
5. EmQue	.01 (–.01)	.16 (.15)	–.29* (–.31)*	.27** (.24*)	–	
6. ET	–.08 (–.11)	–.08 (–.09)	–.13 (–.16)	.17 (.19)	.11 (.11)	–

Note. A-RE = Appearance-Reality Emotion Task (only children who passed both control questions; $n = 50$; values represent point biserial correlations); AST = Affective Stories Task; CSUS = Children's Social Understanding Scale; EmQue = Empathy Questionnaire; ET = Empathy Task; P-RMET = Preschool Reading the Mind in the Eyes Task.

Partial correlations, controlling for age in months, are shown in parentheses. Degrees of freedom ranged from 41 to 116.

* $p < .05$; ** $p < .01$.

age. These findings imply that extant affective ToM tasks share minimal variance beyond that captured by verbal abilities and developmental changes and may not assess a broader affective ToM construct. Instead, the current measures may assess different components of affective ToM or different ways others' affective states can be understood (e.g., by focusing on the eyes and visual cues versus considering the social context, antecedents, and consequences). Thus, caution is encouraged when interpreting findings based on a single affective ToM task, as different measures will lead to different conclusions about the developmental trajectory of affective ToM. Accordingly, we suggest that researchers employ a diverse battery of tasks to capture the multidimensional nature of affective ToM and improve our understanding of children's affective reasoning. Although ToM has been conceptualized in a number of different ways, these differences are rarely discussed when interpreting findings (Schaafsma, Pfaff, Spunt, & Adolphs, 2015). However, more precise accounts of these affective ToM components (e.g., affect recognition from visual cues, and hidden-emotion understanding) and their measurement will clarify inconsistencies and advance the field.

Alternatively, the lack of coherence observed in the current study may reflect the different cognitive demands involved in each affective ToM task. For example, the A-RE Task requires children to maintain two emotional states in mind and, therefore, relies on inhibitory control, working memory, and cognitive flexibility, rather than affective ToM alone. Indeed, inhibition has been found to predict affective ToM performance among adolescents, over and above the effects of age (Vetter, Altgassen, Phillips, Mahy, & Kliegel, 2013). In contrast, performance on the RMET is dependent on one's capacity for affect recognition and/or emotional contagion; thus, the additional influence of these processes may account for the lack of coherence among measures. Although the majority of affective ToM assessments were unrelated after accounting for verbal abilities, it cannot be concluded that the existing relations are driven by verbal abilities alone. Many of the developing cognitive and neural processes that underlie verbal ability also underlie some of the skills needed to complete the affective ToM tasks selected. For example, inhibition and executive functioning have been found to be related to both affective theory of mind and language (e.g., Austin, Groppe, & Elsner, 2014; Carlson & Moses, 2001; Schmitt, Purpura, & Elicker, 2019; Vetter *et al.*, 2013). Thus, verbal ability may reasonably account

for the relations among these variables. Nonetheless, the lack of coherence observed indicates that a single measure should not be regarded as a comprehensive index of affective ToM and that a critical evaluation of existing measures is warranted.

Affective ToM development in preschool

In line with previous research, children's affective ToM increased with age (e.g., Baron-Cohen, O'Riordan, Stone, Jones, & Plaisted, 1999; Wellman & Liu, 2004), such that older children exhibited improved performance and had higher parent ratings of affective ToM relative to younger children. These findings are consistent with Wellman and Liu's (2004) scaling, which indicates that ToM increases throughout preschool until children learn to distinguish between others' real and apparent emotions by the age of six. Similarly, our results are comparable to Pons, Harris, and de Rosnay's (2004) proposed stages of emotion understanding—five-year-olds in the current sample struggled to reconcile hidden emotions, and the majority of children did not succeed at this task until six. Further, performance on the Preschool RMET was consistent across ages, supporting the claim that affect recognition is an early developing aspect of emotion understanding (Pons *et al.*, 2004). Finally, results from the AST revealed significant advancements in children's ability to understand the link between situational causes and expressed emotions (i.e., Component II from Pons *et al.*, 2004) with increasing age.

In the current study, only scores on the AST remained significantly associated with age when children's language abilities were held constant, suggesting that age-related differences on traditional affective ToM tasks may reflect children's developing verbal abilities rather than specific advancements in affective ToM. Indeed, affective ToM tasks with complex affective state terms have commonly been used in ToM research (e.g., Ahmed & Miller, 2011), despite the existence of age-appropriate tasks in the emotion understanding literature (see Harris, 2000 for a review). Moreover, in the current sample, most children did not pass the control questions on the A-RE Task and were omitted from subsequent analyses in accordance with task instructions (Wellman & Liu, 2004). Accordingly, analyses with the A-RE Task were limited in power as they only included a subset of 4-year-olds with acceptable comprehension skills (i.e., 14.6% of 4-year-olds). Further research is needed to determine whether these findings are generalizable beyond the current study, given that they are incompatible with past research (e.g., Gosselin, Madeleine, & Diotte, 2002; Wellman & Liu, 2004). Taken together, the current findings imply that many affective ToM measures require a certain degree of verbal comprehension for children to understand the task instructions and that age-related increases in affective ToM can only be captured, independent of developing verbal abilities, when verbal demands are minimized by using short single-sentence affective stories (i.e., as in the AST).

Affective ToM and empathy

To further examine the convergent validity of affective ToM measures, the current study also investigated the relationship between affective ToM and empathic abilities using an observation-based empathy task and parent reports of empathy. As expected and consistent with previous literature (e.g., Völlm *et al.*, 2006), a positive relationship was found between parent reports of affective ToM and empathy, over and above the effects of age. The lack of associations among other affective ToM and empathy measures, however, suggests that this relationship may be due to shared measurement variance (i.e., parent

report), rather than a conceptual overlap. Further, although a significant negative relation was revealed between the A-RE Task and the EmQue, this unpredicted effect requires replication due to power limitations of our analyses.

Conclusions and implications

This study examined the coherence of child measures of affective ToM and explored the developmental trajectory of affective ToM during preschool. Most affective ToM measures were found to be unrelated, and existing relations were largely driven by the verbal demands of these tasks. Nonetheless, as hypothesized, performance on all measures of affective ToM showed improvements with age and older children were rated by their parents as having superior affective ToM. Further, preliminary support for the utility of the Preschool RMET was found, as scores on this measure showed a positive relation with age. In contrast to our original predictions, however, only the AST captured age-related variance after accounting for verbal abilities, suggesting that traditional measures of affective ToM may rely on children's receptive vocabulary. Given that each of the scales examined showed good variability, with the exception of the A-RE Task, it is unlikely that these findings can be attributed to power limitations. Instead, this lack of coherence suggests that there is a need to examine affective ToM as a theoretical construct and its measurement more extensively.

In particular, we propose that affective ToM be viewed as a collection of independent subcomponents rather than a unitary construct. Indeed, in the emotion understanding literature, emotion understanding is thought to incorporate a number of subprocesses (e.g., affect labelling, facial affect recognition, understanding the antecedents of emotion; Eggum *et al.*, 2011); however, ToM is often not conceptualized in this way. Future work should examine the relations among conceptually similar affective ToM tasks to capture its multidimensional nature. Next steps could assess the coherence among visual affect recognition tasks or those that require children to identify others' emotions based on stories/vignettes. Using a combination of behavioural tasks and neuroimaging techniques will help to validate these conceptual distinctions by examining whether identified subcomponents correspond to distinct neural underpinnings. Given the impact of verbal abilities on children's affective ToM performance, we suggest that researchers use tasks with minimal verbal demands or control for verbal ability in affective ToM performance. Once dissociable subcomponents of affective ToM are identified (e.g., affect recognition, and hidden emotion), they can be used to form a multidimensional affective ToM battery that may better capture the complexity of this construct, predict social outcomes (e.g., social participation and friendships), and account for patterns of strengths and weaknesses across subcomponents in clinical populations.

Acknowledgements

CG received support from the Vanier Canada Graduate Scholarships (Vanier CGS) program during the preparation, data collection, analysis, and writing of this research. CEVM received support from an NSERC Discovery Grant (RGPIN-2015-03774). The authors wish to thank Tessa Mazachowsky, Chelsey Masson, Katarina McKenzie, and Amanda Krause for their help with data collection.

Conflicts of interest

All authors declare no conflict of interest.

Author contributions

Caitlyn Gallant (Conceptualization; Formal analysis; Investigation; Methodology; Project administration; Supervision; Validation; Visualization; Writing – original draft; Writing – review & editing) Lydia Lavis (Project administration; Writing – review & editing) Caitlin E. V. Mahy (Conceptualization; Formal analysis; Funding acquisition; Investigation; Methodology; Project administration; Resources; Supervision; Validation; Visualization; Writing – original draft; Writing – review & editing).

Data availability statement

The data that support the findings of this study are available on request from the corresponding author. The data are not publicly available due to privacy or ethical restrictions.

References

- Ahmed, F. S., & Miller, S. L. (2011). Executive function mechanisms of theory of mind. *Journal of Autism and Developmental Disorders*, *41*(5), 667–678. <https://doi.org/10.1007/s10803-010-1087-7>
- Austin, G., Groppe, K., & Elsner, B. (2014). The reciprocal relationship between executive function and theory of mind in middle childhood: A 1-year longitudinal perspective. *Frontiers in Psychology*, *5*, 655. <https://doi.org/10.3389/fpsyg.2014.00655>
- Baron-Cohen, S., O'Riordan, M., Stone, V., Jones, R., & Plaisted, K. (1999). Recognition of faux pas by normally developing children and children with Asperger syndrome or high-functioning autism. *Journal of Autism and Developmental Disorders*, *29*, 407–418. <https://doi.org/10.1023/A:1023035012436>
- Baron-Cohen, S., Tager-Flusberg, H., & Cohen, D. J. (Eds.) (1994). *Understanding other minds: Perspectives from autism*. New York, NY: Oxford University Press.
- Baron-Cohen, S., Wheelwright, S., Spong, A., Scahill, V., & Lawson, J. (2001). Are intuitive physics and intuitive psychology independent? A test with children with Asperger Syndrome. *Journal of Developmental and Learning Disorders*, *5*, 47–78.
- Blair, C., & Razza, R. P. (2007). Relating effortful control, executive function, and false belief understanding to emerging math and literacy ability in kindergarten. *Child Development*, *78*, 647–663. <https://doi.org/10.1111/j.1467-8624.2007.01019.x>
- Butean, I., Costescu, C., & Dobrean, A. (2014). Differences between empathic responses in children with autism spectrum disorder and typically developing children. *Journal of Evidence-Based Psychotherapies*, *14*(2), 197–209.
- Caputi, M., Lecce, S., Pagnin, A., & Banerjee, R. (2012). Longitudinal effects of theory of mind on later peer relations: The role of prosocial behavior. *Developmental Psychology*, *48*(1), 257–270. <https://doi.org/10.1037/a0025402>
- Carlson, S. M., & Moses, L. J. (2001). Individual differences in inhibitory control and children's theory of mind. *Child Development*, *72*, 1032–1053. <https://doi.org/10.1111/1467-8624.00333>
- Carroll, J. M., & Chiew, K. Y. (2006). Sex and discipline differences in empathising, systemising and autistic symptomatology: Evidence from a student population. *Journal of Autism and Developmental Disorders*, *36*, 949–957. <https://doi.org/10.1007/s10803-006-0127-9>

- Castelli, F., Frith, C., Happé, F., & Frith, U. (2002). Autism, Asperger syndrome and brain mechanisms for the attribution of mental states to animated shapes. *Brain*, *125*(8), 1839–1849. <https://doi.org/10.1093/brain/awf189>
- Cutting, A. L., & Dunn, J. (1999). Theory of mind, emotion understanding, language, and family background: Individual differences and interrelations. *Child Development*, *70*, 853–865. <https://doi.org/10.1111/1467-8624.00061>
- Denham, S. A. (1986). Social cognition, prosocial behavior, and emotion in preschoolers: Contextual validation. *Child Development*, *194–201*, <https://doi.org/10.2307/1130651>
- Egger, H. L., Pine, D. S., Nelson, E., Leibenluft, E., Ernst, M., Towbin, K. E., & Angold, A. (2011). The NIMH Child Emotional Faces Picture Set (NIMH-ChEFS): A new set of children's facial emotion stimuli. *International Journal of Methods in Psychiatric Research*, *20*(3), 145–156. <https://doi.org/10.1002/mpr.343>
- Eggum, N. D., Eisenberg, N., Kao, K., Spinrad, T. L., Bolnick, R., Hofer, C., . . . Fabricius, W. V. (2011). Emotion understanding, theory of mind, and prosocial orientation: Relations over time in early childhood. *The Journal of Positive Psychology*, *6*(1), 4–16. <https://doi.org/10.1080/17439760.2010.536776>
- Eisenberg, N., Spinrad, T. L., & Sadovsky, A. (2006). Empathy-related responding in children. In M. Killen & J. Smetana (Eds.), *Handbook of moral development* (pp. 517–549). Mahwah, NJ: Erlbaum.
- Ekman, P., & Friesen, W. (1975). *Unmasking the face*. Englewood Cliffs, NJ: Prentice-Hall.
- Gallese, V., & Goldman, A. (1998). Mirror neurons and the simulation theory of mind-reading. *Trends in Cognitive Sciences*, *2*, 493–501. [https://doi.org/10.1016/S1364-6613\(98\)01262-5](https://doi.org/10.1016/S1364-6613(98)01262-5)
- Garner, P. W., Jones, D. C., & Miner, J. L. (1994). Social competence among low-income preschoolers: Emotion socialization practices and social cognitive correlates. *Child Development*, *65*, 622–637. <https://doi.org/10.1111/j.1467-8624.1994.tb00772.x>
- Ginsburg, H. J., Ogletree, S. M., Silakowski, T. D., Bartels, R. D., Burk, S. L., & Turner, G. M. (2003). Young children's theories of mind about empathic and selfish motives. *Social Behavior and Personality: An International Journal*, *31*, 237–243. <https://doi.org/10.2224/sbp.2003.31.3.237>
- Gopnik, A. (2003). The theory theory as an alternative to the innateness hypothesis. In L. M. Antony & N. Hornstein (Eds.), *Chomsky and his critics* (pp. 238–254). Oxford, UK: Blackwell.
- Gosselin, P., Warren, M., & Diotte, M. (2002). Motivation to hide emotion and children's understanding of the distinction between real and apparent emotions. *The Journal of Genetic Psychology*, *163*(4), 479–495. <https://doi.org/10.1080/00221320209598697>
- Harris, P. L. (1989). *Children and emotion: The development of psychological understanding*. Cambridge, MA: Basil Blackwell.
- Harris, P. L. (1992). From simulation to folk psychology: The case for development. *Mind and Language*, *7*, 120–144. <https://doi.org/10.1111/j.1468-0017.1992.tb00201.x>
- Harris, P. L. (2000). Understanding emotion. In M. Lewis & J. M. Haviland-Jones (Eds.), *Handbook of emotions* (pp. 281–292). New York, NY: Guilford.
- Kalbe, E., Grabenhorst, F., Brand, M., Kessler, J., Hilker, R., & Markowitsch, H. J. (2007). Elevated emotional reactivity in affective but not cognitive components of theory of mind: A psychophysiological study. *Journal of Neuropsychology*, *1*, 27–38. <https://doi.org/10.1348/174866407X180792>
- Kalbe, E., Schlegel, M., Sack, A. T., Nowak, D. A., Dafotakis, M., Bangard, C., . . . Kessler, J. (2010). Dissociating cognitive from affective theory of mind: A TMS study. *Cortex*, *46*, 769–780. <https://doi.org/10.1016/j.cortex.2009.07.010>
- Klin, A., Jones, W., Schultz, R., Volkmar, F., & Cohen, D. (2002). Visual fixation patterns during viewing of naturalistic social situations as predictors of social competence in individuals with autism. *Archives of General Psychiatry*, *59*(9), 809–816. <https://doi.org/10.1001/archpsyc.59.9.809>

- Lawrence, E. J., Shaw, P., Baker, D., Baron-Cohen, S., & David, A. S. (2004). Measuring empathy: Reliability and validity of the Empathy Quotient. *Psychological Medicine*, *34*, 911–920. <https://doi.org/10.1017/S0033291703001624>
- Oakley, B. F., Brewer, R., Bird, G., & Catmur, C. (2016). Theory of mind is not theory of emotion: A cautionary note on the Reading the Mind in the Eyes Test. *Journal of Abnormal Psychology*, *125* (6), 818. <https://doi.org/10.1037/abn0000182>
- Peterson, R. A. (1994). A meta-analysis of Cronbach's coefficient alpha. *Journal of Consumer Research*, *21*(2), 381–391. <https://doi.org/10.1086/209405>
- Pons, F., & Harris, P. (2000). *Test of emotion comprehension: TEC*. Oxford, UK: University of Oxford.
- Pons, F., Harris, P. L., & de Rosnay, M. (2004). Emotion comprehension between 3 and 11 years: Developmental periods and hierarchical organization. *European Journal of Developmental Psychology*, *1*(2), 127–152. <https://doi.org/10.1080/17405620344000022>
- Rieffe, C., Ketelaar, L., & Wieferrink, C. H. (2010). Assessing empathy in young children: Construction and validation of an Empathy Questionnaire (EmQue). *Personality and Individual Differences*, *49*, 362–367. <https://doi.org/10.1016/j.paid.2010.03.046>
- Schaafsma, S. M., Pfaff, D. W., Spunt, R. P., & Adolphs, R. (2015). Deconstructing and reconstructing theory of mind. *Trends in Cognitive Sciences*, *19*, 65–72. <https://doi.org/10.1016/j.tics.2014.11.007>
- Schmitt, S. A., Purpura, D. J., & Elicker, J. G. (2019). Predictive links among vocabulary, mathematical language, and executive functioning in preschoolers. *Journal of Experimental Child Psychology*, *180*, 55–68. <https://doi.org/10.1016/j.jecp.2018.12.005>
- Sebastian, C. L., Fontaine, N. M., Bird, G., Blakemore, S. J., De Brito, S. A., McCrory, E. J., & Viding, E. (2011). Neural processing associated with cognitive and affective theory of mind in adolescents and adults. *Social Cognitive and Affective Neuroscience*, *7*(1), 53–63. <https://doi.org/10.1093/scan/nsr023>
- Shamay-Tsoory, S. G., Tomer, R., Berger, B. D., Goldsher, D., & Aharon-Peretz, J. (2005). Impaired “affective theory of mind” is associated with right ventromedial prefrontal damage. *Cognitive and Behavioral Neurology*, *18*, 55–67. <https://doi.org/10.1097/01.wnn.0000152228.90129.99>
- Shamay-Tsoory, S. G., Tomer, R., Goldsher, D., Berger, B. D., & Aharon-Peretz, J. (2004). Impairment in cognitive and affective empathy in patients with brain lesions: Anatomical and cognitive correlates. *Journal of Clinical and Experimental Neuropsychology*, *26*, 1113–1127. <https://doi.org/10.1080/13803390490515531>
- Slaughter, V., Dennis, M. J., & Pritchard, M. (2002). Theory of mind and peer acceptance in preschool children. *British Journal of Developmental Psychology*, *20*, 545–564. <https://doi.org/10.1348/026151002760390945>
- Stanford, A. D., Messinger, J., Malaspina, D., & Corcoran, C. M. (2011). Theory of Mind in patients at clinical high risk for psychosis. *Schizophrenia Research*, *131*, 11–17. <https://doi.org/10.1016/j.schres.2011.06.005>
- Strayer, J., & Roberts, W. (1989). Children's empathy and role taking: Child and parental factors, and relations to prosocial behavior. *Journal of Applied Developmental Psychology*, *10*, 227–239. [https://doi.org/10.1016/0193-3973\(89\)90006-3](https://doi.org/10.1016/0193-3973(89)90006-3)
- Tahiroglu, D., Moses, L. J., Carlson, S. M., Mahy, C. E. V., Olofson, E. L., & Sabbagh, M. A. (2014). The Children's Social Understanding Scale: Construction and validation of a parent-report measure for assessing individual differences in children's theories of mind. *Developmental Psychology*, *50*(11), 2485–2497. <https://doi.org/10.1037/a0037914>
- Vetter, N. C., Altgassen, M., Phillips, L., Mahy, C. E., & Kliegel, M. (2013). Development of affective theory of mind across adolescence: Disentangling the role of executive functions. *Developmental Neuropsychology*, *38*(2), 114–125. <https://doi.org/10.1080/87565641.2012.733786>
- Völlm, B. A., Taylor, A. N., Richardson, P., Corcoran, R., Stirling, J., McKie, S., . . . Elliott, R. (2006). Neuronal correlates of theory of mind and empathy: A functional magnetic resonance imaging

- study in a nonverbal task. *NeuroImage*, 29, 90–98. <https://doi.org/10.1016/j.neuroimage.2005.07.022>
- Warnell, K. R., & Redcay, E. (2019). Minimal coherence among varied theory of mind measures in childhood and adulthood. *Cognition*, 191, 103997. <https://doi.org/10.1016/j.cognition.2019.06.009>
- Wellman, H. M., & Liu, D. (2004). Scaling of theory-of-mind tasks. *Child Development*, 75, 523–541. <https://doi.org/10.1111/j.1467-8624.2004.00691.x>
- Zahn-Waxler, C., Robinson, J. L., & Emde, R. N. (1992). The development of empathy in twins. *Developmental Psychology*, 28(6), 1038–1047. <https://doi.org/10.1037/0012-1649.28.6.1038>
- Zelazo, P. D., Bauer, P. J., (Eds.) (2013). *National Institutes of Health Toolbox cognition battery (NIH Toolbox CB): Validation for children between 3 and 15 years*. Hoboken, NJ : Wiley.

Received 14 September 2019; revised version received 5 March 2020