Journal of Experimental Child Psychology 149 (2016) 98–115

Contents lists available at ScienceDirect

Journal of Experimental Child Psychology

journal homepage: www.elsevier.com/locate/jecp

Theory of mind in middle childhood and early adolescence: Different from before?

Nancie Im-Bolter [a,*], Alba Agostino [b], Keely Owens-Jaffray [a,c]

[a] Department of Psychology, Trent University, Peterborough, Ontario K9L 0G2, Canada
[b] Department of Psychology, Ryerson University, Toronto, Ontario M5B 2K3, Canada
[c] Department of Psychology, Queen's University, Kingston, Ontario K7L 3N6, Canada

ARTICLE INFO

Article history:
Available online 27 January 2016

Keywords:
Theory of mind
Mental attentional capacity
Inhibition
Shifting
Updating
Language

ABSTRACT

Studies with preschool children have shown that language and executive function are important for theory of mind, but few studies have examined these associations in older children and in an integrative theory-guided manner. The theory of constructive operators was used as a framework to test a model of relations among mental attentional capacity, attentional inhibition, language, executive processes (shifting and updating), and higher order theory of mind in two groups of school-aged children: one in middle childhood ($n = 226$; mean age = 8.08 years) and the other in early adolescence ($n = 216$; mean age = 12.09 years). Results revealed a complex model of interrelations between cognitive resources and language in middle childhood that directly and indirectly predicted theory of mind. The model in early adolescence was less complex, however, and highlighted the importance of semantic language and shifting for theory of mind. Our findings suggest not only that contributors to theory of mind change over time but also that they may depend on the maturity level of the theory of mind system being examined.

Introduction

The majority of research in theory of mind (ToM), or understanding of mental perspective (e.g., beliefs, intents), has focused on the preschool period, although advancements in ToM are believed

* Corresponding author.
E-mail address: nimbolter@trentu.ca (N. Im-Bolter).

http://dx.doi.org/10.1016/j.jecp.2015.12.006

to continue into middle childhood and adolescence (Miller, 2009). In the preschool literature, both language and executive function are related to ToM (Astington & Jenkins, 1999; Carlson & Moses, 2001; Devine & Hughes, 2014; Milligan, Astington, & Dack, 2007). The limited research with older children generally is consistent with these findings (see Miller, 2009, for a review); however, no study to our knowledge has examined an integrative model of the relations among language, executive function, and ToM in middle childhood and early adolescence in a theory-guided manner. In the current study, we test two models of relations among mental attentional (*M*) capacity, attentional inhibition (*I*), language, and executive function within a comprehensive theoretical framework (Pascual-Leone's theory of constructive operators) to examine how these processes contribute to ToM in middle childhood compared with early adolescence.

Theory of mind: Preschool children

Critical developments in children's understanding of mental states occur during a period when children are acquiring more sophisticated structural (semantic and syntactic) language skills. There is now a well-established literature regarding the importance of semantic and syntactic language to ToM in both typical (e.g., Astington & Jenkins, 1999; Milligan et al., 2007) and atypical populations (e.g., autism spectrum disorder: Happé, 1994; Tager-Flusberg & Joseph, 2005; language impairment: Farrant, Fletcher, & Maybery, 2006; Gillott, Furniss, & Walter, 2004). This makes sense when we consider language is the primary means by which we communicate and acquire knowledge about the mental world. Beliefs and intentions are physically unobservable, as is their relation to behavior. Semantics provide a means for representing unobservable mental states (e.g., *think*, *know*, *believe*), and syntax provides a structure for representing and keeping track of false beliefs (e.g., Mary *thinks the doll is in the box*) as well as reflecting on self and other beliefs (e.g., I *thought* she *knew* he was going).

Although structural language clearly plays a role in ToM, the research regarding the relative importance of semantic versus syntactic language is equivocal. Some studies provide support for the primary role of semantics (Markel, Major, & Pelletier, 2013), and some provide support for the primary role of syntax (Astington & Jenkins, 1999; de Villiers & Pyers, 2002). Others have argued that semantic and syntactic language are too highly correlated and cannot be disentangled from each other (Ruffman, Slade, Rowlandson, Rumsey, & Garnham, 2003). A meta-analysis by Milligan and colleagues (2007) suggests that although structural language accounts for an impressive amount of variance in false belief task performance (used to measure first-order ToM), the strength of this relation is quite variable (from small negative to large effect sizes). This raises the question of other cognitive factors that might contribute to ToM.

A highly researched correlate of ToM is executive function (EF), an umbrella term used to describe distinct but related abilities that direct, organize, and mediate problem solving. A three-factor model of EF (inhibition, updating of working memory contents, and shifting of mental sets), originally found in adults (Miyake et al., 2000), has been replicated in studies with school-aged children (Lehto, Juujarvi, Kooistra, & Pulkkinen, 2003; Rose, Feldman, & Jankowski, 2011). The structure of EF is less clear in younger children, with empirical evidence for both a unitary model (Brydges, Reid, Fox, & Anderson, 2012) and a two-factor model (Miller, Giesbrecht, Müller, McInerney, & Kearns, 2012; Miyake & Friedman, 2012). Regardless of the structure, EF has been shown to be associated with first-order ToM in preschoolers (Carlson & Moses, 2001; Devine & Hughes, 2014; Perner & Lang, 1999), particularly inhibition (Carlson, Mandell, & Williams, 2004; Carlson, Moses, & Breton, 2002) and shifting or cognitive flexibility (Farrant, Maybery, & Fletcher, 2012; Low, 2010). Theoretically, EF would assist in distinguishing, coordinating, and tracking different mental intentions. A meta-analysis by Devine and Hughes (2014) showed a moderate association between EF and false belief understanding (15% shared variance), which remained significant (8% shared variance) after accounting for verbal ability. It should be highlighted, however, that verbal ability might be measured by a single vocabulary test (e.g., Carlson & Moses, 2001), so these findings do not help to clarify the relations among language, EF, and ToM.

Given the importance of language and EF to ToM, there is a surprising lack of studies integrating these two areas of research in younger children. Studies conducted so far (Benson, Sabbagh, Carlson, & Zelazo, 2013; Farrant et al., 2012; Hughes, 1998; Low, 2010; Müller, Liebermann-Finestone, Carpendale, Hammond, & Bibok, 2012) show a contradictory picture of the relations among language, EF, and ToM. After accounting for verbal ability, EF has been found to be both related and not related to false belief understanding (cf. Benson et al., 2013, and Hughes, 1998). Yet other studies show that both language and EF (shifting) are important for false belief understanding in 3- and 4-year-olds (Low, 2010) and 3- to 6-year-olds (Farrant et al., 2012). Interestingly, Müller and colleagues (2012) found that after accounting for verbal ability, EF was related to false belief understanding in 3- and 4-year-olds but not in 2-year-olds, suggesting involvement of different processes at different ages. Nonetheless, there remains little clarity regarding the specific relations among language, EF, and ToM in younger children and whether these relations continue as children get older.

Theory of mind: Beyond the preschool years

Little is known about the relations among language, EF, and ToM beyond the preschool period even though there is a growing literature showing that ToM continues to develop in complexity (Apperly, Warren, Andrews, Grant, & Todd, 2011; Devine & Hughes, 2013; Dumontheil, Apperly, & Blakemore, 2010; Vetter, Leipold, Kliegel, Phillips, & Altgassen, 2013; see also Apperly, Samson, & Humphreys, 2009, and Miller, 2009, for reviews). In older children (and adults), measurement of advanced or higher order ToM often involves a recursive or hierarchical element of mental state understanding (i.e., thinking about another's thinking of another) or attribution of intentional mental states to explain others' behavior. Higher order ToM tasks have been developed to assess understanding of others' intentions/beliefs (e.g., Happé, 1994; Vetter, Leipold et al., 2013) and visual perspectives (e.g., Dumontheil et al., 2010; Keysar, Lin, & Barr, 2003). Research with these tasks indicates that higher order ToM develops with age (Devine & Hughes, 2013; Dumontheil et al., 2010; Keysar et al., 2003) and that improvements in higher order ToM continue to occur between adolescence and adulthood (Dumontheil et al., 2010; Vetter, Leipold et al., 2013). The literature suggests that even adults show limits in their ability to use their perspective-taking abilities to make inferences about others' behavior (Apperly et al., 2010, 2011; Keysar et al., 2003). These results have led to the proposal that language and EF may be important for developing ToM in the preschool period, but this may change as the ToM system matures (Apperly et al., 2009).

The relative importance of language and EF to higher order ToM in older children is not clear. Research conducted so far suggests that both language (Bosacki, 2000; Gillott et al., 2004; Miller, 2001) and EF (Ahmed & Miller, 2011; Austin, Gropper, & Elsner, 2014; Vetter, Altgassen, Phillips, Mahy, & Kliegel, 2013) continue to play a role in higher order ToM. Interestingly, consistent with what has been found with preschoolers, research by Ahmed and Miller (2011) suggests that shifting or cognitive flexibility may be particularly important for higher order theory of mind in adults. There appear to be no studies, however, that have examined the simultaneous contribution of language and EF to higher order ToM in older children. Moreover, to our knowledge, the current study is the first to investigate these relations in a theory-guided manner to determine whether, and how, these relations change over development.

Theory of constructive operators

The theory of constructive operators (Pascual-Leone, 1970, 1987; Pascual-Leone & Johnson, 2005, 2011) views cognitive growth in terms of the maturation of domain-general central processing resources—mental attentional activation (*M*) capacity and attentional inhibition (interruption or *I*) capacity—as well as mechanisms for logical-structural and content learning. In the theory of constructive operators, mental attention is explained by activatory (*M*), inhibitory (*I*), and executive (*E*) processes (as well as other constructs; see Pascual-Leone & Johnson, 2011, for a detailed discussion). *M* raises the activation level of task-relevant schemes (operative processes and mental representations) that are not sufficiently activated by the situation, *I* inhibits or effortfully lowers activation of task-irrelevant schemes, and *E* controls and monitors allocation of *M* and *I* and has a general higher order planning function. When assessed behaviorally, the capacity of *M* is proposed to increase by 1 scheme unit every other year from 3 to 15 years of age (e.g., 7- and 8-year-olds have an *M* capacity of 3, and

11- and 12-year-olds have an *M* capacity of 5). Research with measures developed to evaluate *M* capacity in various domains, including visual spatial and verbal, supports this proposed growth of *M* capacity (Arsalidou, Pascual-Leone, & Johnson, 2010; Pascual-Leone & Johnson, 1999, 2005; Powell, Arsalidou, Vogan, & Taylor, 2014).

It is possible to bypass irrelevant schemes via activation rather than with *I* (Pascual-Leone, 1987). When a strategy effectively excludes misleading schemes, the act of focal attention will suppress them automatically. This, however, may require additional *M* because often more schemes must be activated to apply this new detour/bypass strategy. Theoretically, this suggests that young children, who have less *M*, would have difficulty with tasks that involve misleading schemes because they might not mobilize *I* efficiently and might be unable to use detour/bypass strategies. Older children, who have more *M*, might bypass irrelevant schemes or effortfully inhibit them. This dual option of a detour/bypass strategy versus effortful inhibition of task-misleading schemes clarifies the sense in which inhibition is related to controlled attentional resources (Engle, Conway, Tuholski, & Shisler, 1995). Note that, unlike some approaches (e.g., Baddeley, 1996; Miyake et al., 2000), the theory of constructive operators considers effortful (direct) inhibition to be a cognitive resource, not an EF, although executive processes are used to control it (Agostino, Johnson, & Pascual-Leone, 2010; Howard, Johnson, & Pascual-Leone, 2014; Im-Bolter, Johnson, & Pascual-Leone, 2006). *E* (executive processes) monitors allocation of *M* and *I* to serve the current goal. Problem-solving situations vary in their demand for *M* and executive control. Thus, the efficiency with which an individual can mobilize and allocate *M* and *I* will depend on *E* as well as characteristics of the problem-solving situation, previous experiences, and the person's repertoire of control executives.

Particularly relevant to the current study are the *M* control executives recentration and decentration. *Recentration* changes the content of focal attention without shifting levels of analysis. An example is scanning a room or, in language tasks, from schemes (visuals patterns or words) already cognized to new schemes just activated by input. In contrast, *decentration* monitors or controls shifting of focal attention to schemes constituted at a higher or lower level of analysis. For example, when one needs to understand a complex and unfamiliar language utterance, the units already cognized and the new units being attended to must be synthesized into a hierarchically organized totality. This meaningful totality integrates parts into a composite, higher order meaning structure. Recentration and decentration correspond most closely to updating and shifting, respectively, as formulated by Miyake and colleagues (2000) in their conceptualization of EF. Note that in the theory of constructive operators, recentration and decentration provide executive control during the allocation of *M*, which is distinct from *M* activation processes.

The current study

Using the theory of constructive operators, we tested distinct theoretical models of higher order ToM (see Fig. 1) in two groups of children with unequivocal differences in their theoretical *M* capacity: (a) 7- and 8-year-olds with an *M* of 3 (middle childhood) and (b) 11- and 12-year-olds with an *M* of 5 (early adolescence). According to the theory of constructive operators, the models are different in these age groups because of two important changes that occur during this time: (a) an increase in *M* capacity of 2 scheme units between the younger and older groups, which allows the older group to keep active more relevant information, and (b) an increase in experience with the higher order ToM system (in terms of both the processes and contexts involved). These differences change the cognitive demand for higher order ToM experienced by each age group and are reflected in the two models. According to the theory of constructive operators, characteristics of a problem-solving situation will place different demands on *M*, *I*, and *E*. In a typical higher order ToM task that assesses the ability to attribute intentional mental states, the individual must (a) keep active multiple units of relevant information (with *M*), (b) change the content of focal attention with each incoming unit of relevant information (with *recentration*) so that each individual's intentional mental state is kept up to date, (c) shift the content of focal attention (with *decentration*) between the intentional mental states of each individual in order to explain the resulting behavior, (d) inhibit irrelevant or misleading information (with *I*) in order to accurately represent each individual's mental intentions, (e) use semantic language to represent mental states and emotions as well as modified, elaborated, contrasting, or

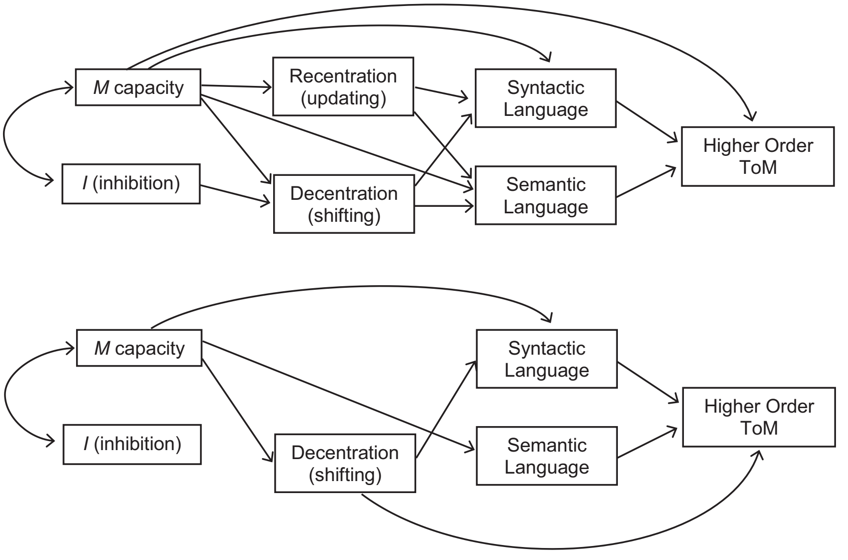

Fig. 1. Proposed theoretical model of the relations among *M*, *I*, recentration, decentration, syntactic and semantic language, and higher order theory of mind. Top panel: Middle childhood. Bottom panel: Early adolescence.

dependent mental states signaled by conjunctions (e.g., *though*, *because*, *but*, *when*), and (f) use syntactic language to represent, scaffold, and keep track of intentions and beliefs that may be hierarchically embedded (e.g., "Mrs. Smith knows Jill is not sure about buying a kitten but is trying to make her want one").

Younger school-aged children can be considered novices with respect to higher order ToM (similar to preschoolers and first-order ToM). Therefore, the middle childhood model reflects the more intensive involvement of all cognitive resources, particularly *M*, which is needed to keep active and hold in mind all relevant information. In this age group, the model reflects that ToM is predicted by *M* capacity with executive processes modulating its function. Based on the model of language competence found by Im-Bolter and colleagues (2006), which indicates that recentration and decentration mediate the relation between *M* capacity and language, we hypothesize that the relation between executive processes and ToM is through language. Within the theory of constructive operators, inhibition (*I*) is not an EF but rather an attentional resource that modulates application of *M*; therefore, the use of *I* is associated with the controlled use of *M*. As a result, *I* has an interactive relation with *M* but not language or ToM. Previous research in the domain of language confirms this association (Im-Bolter, Johnson, Ling, & Pascual-Leone, 2015), which we also apply to ToM. Finally, we include a path from *I* to decentration because the application of *I* assists in the shifting of focal attention from one perspective to another. Note that younger children have reduced *M* capacity and, in a situation that is demanding of *M* resources, are not able to use *M* to boost schemes that will allow a detour/bypass strategy. This necessitates the use of decentration to control the use of *I*. The top panel in Fig. 1 illustrates the proposed theoretical model to be tested for middle childhood.

The model of higher order ToM for early adolescence reflects developmental gains in *M* acquired by this age group and their increased experience with mental perspectives. As a result, this model shows less involvement of all cognitive resources, particularly *M*, and fewer paths compared with the middle childhood model. Older children are able to use a detour/bypass strategy where additional *M* can be applied instead of *I*. As a result, there is no path from *I* to decentration, and the contribution of *M* for higher order ToM is based on its association with language and EF only. Similarly, recentration has been removed from the model because increased *M* allows the content of focal attention to be maintained and flexibly shifted from one perspective to the other. For this reason, we propose a path from

M to decentration (shifting) to higher order ToM, which highlights the increased importance of decentration in this age group. There is no path from decentration to semantic language because higher order ToM no longer requires the synthesization of unfamiliar meanings into a hierarchically organized whole. The bottom panel in Fig. 1 illustrates the proposed theoretical model to be tested for early adolescence.

Method

Participants

Recruitment

Participants came from 14 schools just outside a large metropolitan Canadian city. The schools were in generally middle-class suburban neighborhoods, and participants included children from different ethnic backgrounds (e.g., families of Southeast Asian and Asian descent), although the majority of the sample was Caucasian. A total of 228 children aged 7 or 8 years (mean age = 8.08 years, *SD* = 0.51) were included in the middle childhood group. This group consisted of 121 boys and 107 girls. A total of 216 children aged 11 or 12 years (mean age = 12.09 years, *SD* = 0.49) were included in the early adolescence group. This group consisted of 106 boys and 110 girls. Parental consent and child assent were obtained for all children.

Inclusion and exclusion criteria

All children met the following criteria: (a) estimated Performance IQ score within the average range (i.e., 80–129 on the Wechsler Abbreviated Scale of Intelligence–Matrix Analogies subtest), (b) English spoken in the home without significant dialectical differences, (c) not attending an English-as-a-second-language or -dialect program, (d) no hard signs of neurological damage and not diagnosed with a developmental disorder, (e) no diagnosis of a behavioral or psychiatric disorder, and (f) not receiving any services (formally or informally) for any language, learning, behavioral, or emotional problems and no history of problems in these areas based on interviews with school personnel (e.g., special education resource teacher).

Measures

Socioeconomic status

The Blishen Scale (Blishen, Carroll, & Moore, 1987) was used to code socioeconomic status (SES). This scale uses a combination of education level and occupation to determine level of SES and ranges from 17.81 (newspaper carriers and vendors) to 101 (physicians and surgeons).

Nonverbal intelligence

The Matrix Analogies subtest of the Wechsler Abbreviated Scale of Intelligence (WASI; Wechsler, 1999) was used to estimate Performance IQ (PIQ).

Language skills

Each child received a short battery of standardized tests compiled to measure areas typically assessed by speech/language pathologists. These included the expressive and receptive components of semantics and syntax. The Peabody Picture Vocabulary Test–Fourth Edition (PPVT-4; Dunn & Dunn, 2006) measured *receptive semantics*, and the Expressive Vocabulary Test–Second Edition (EVT-2; Williams, 2006) measured *expressive semantics*. The Test of the Reception of Grammar–Second Edition (TROG-2; Bishop, 2003) assessed *receptive syntax*, and the Clinical Evaluation Language Fundamentals–Fourth Edition–Formulated Sentences subtest (CELF-4-FS; Semel, Wiig, & Secord, 2003) assessed *expressive syntax*. The standard scores from the language tests were transformed to z scores to create a composite Semantic language score (mean of the PPVT-4 and EVT-2 z scores) and a composite Syntactic language score (mean of the TROG-2 and CELF-4-FS z scores).

M capacity

Two individually administered *M* capacity measures (*M* measures), the *direction following task* (DFT) and the *figural intersections test* (FIT), were used to measure *verbal M capacity* and *visuospatial M capacity*, respectively. Test administration followed a standard protocol, and training items were given to provide practice and feedback and to ensure that each child was sufficiently familiar with the task and instructions. Construct validity and developmental data across specific domains (e.g., language, visuospatial) and cross-culturally show that the DFT and FIT contain items that vary systematically in their demand for *M* capacity or *M* demand (Miller, Pascual-Leone, Campbell, & Juckes, 1989; Pascual-Leone & Ijaz, 1989; Pascual-Leone & Johnson, 2005). The test items differ only in terms of the number of schemes that must be kept simultaneously activated by *M* capacity (i.e., *M* demand). As a result, the DFT and FIT have the same scaling and yield the same metric across the verbal and visuospatial domains. This allows direct comparison of performance between the two measures. For example, an *M* score of 3 on the DFT is comparable to an *M* score of 3 on the FIT. Working memory span tasks have not been developed in a similar manner; for this reason, these measures are not just working memory span tasks.

The DFT and FIT exhibit high reliability and validity and are good predictors of performance on other cognitive tasks (Agostino et al., 2010; Im-Bolter et al., 2006; Johnson, Im-Bolter, & Pascual-Leone, 2003; Pascual-Leone & Johnson, 1999, 2005). Each *M* measure provided an *M* score; this corresponded to the *M* demand of the highest performance level that the child consistently attained on a task.

The DFT required children to follow oral directions of increasing complexity (Cunning, 2003; Pascual-Leone & Johnson, 2005). The task employed tokens of basic shapes, colors, and sizes as well as a simple repetitive command ("place X on Y") to control for extraneous factors (e.g., preposition difficulty, degree of abstractness) that load heavily on experiential learning. Children placed tokens on spaces that varied in color and size. There were five items at each of seven levels of complexity (e.g., Level 1: Place a blue square on a white space; Level 4: Place a large square and a small circle on a large space). Items were presented in order of increasing complexity.

The FIT (Version 8303) (Pascual-Leone & Ijaz, 1989; Pascual-Leone & Johnson, 2001) required children to locate the one area of intersection of two to eight overlapping geometric shapes. Children first placed a dot in each discrete shape on the right side of the page and then placed a single dot in the intersection area of the overlapping configuration on the left. There were 36 items, and all participants received the same random order of items. A composite *M* score was created by taking the average of the DFT *M* score and the FIT *M* score.

Attentional inhibition (I)

The *antisaccade task* (Im-Bolter et al., 2006; Miyake et al., 2000), a computer-based task, was used to measure *I* (inhibition). Children sat approximately 18 inches from, and with eye level at the vertical middle of, the computer screen. All stimuli were presented at fixation. The antisaccade task has been found to be sensitive to prefrontal dysfunction and is widely viewed as a task of inhibitory control (see Roberts, Hager, & Heron, 1994, for a short review). It has been used with a wide variety of populations, including children, because it is simple and nonverbal and has minimal memory demands; at the same time, adults do not perform at ceiling levels (Roberts et al., 1994). Children were presented with a fixation point at the center of the computer screen for an amount of time that randomly varied between 1500 and 3500 ms, followed by a blank screen for 50 ms and then a visual cue (small black square) on one side of the screen (e.g., left) for 225 ms. A blank screen followed for 50 ms, and then a target stimulus appeared on the opposite side of the screen (e.g., right) for 100 ms and was then masked. The target stimulus consisted of a light gray arrow (pointing left, right, or up) inside a square. Children indicated the direction of the arrow (left, right, or up) with a button press response. To see the direction of the arrow, they needed to inhibit the reflexive response of looking at the initial visual cue because the target stimulus appeared on the screen for a very brief moment before being masked. Children received 22 practice trials and 90 target trials. Cue location and arrow direction were counterbalanced across trials and were presented in an individually determined random order. The computer recorded response latency (in milliseconds) and accuracy of responses. The score was the proportion of correct responses.

Children received a response mapping task before the actual antisaccade task. The response mapping task gave them proficiency with the response box and ensured that all children had memorized the button press response for the left, right, and up arrows. Presentation of fixation was the same as in the antisaccade task. After presentation of fixation, there was a 100-ms pause and then a light gray arrow appeared in the center of the screen for 100 ms; this was followed by a mask. Children were instructed to make the appropriate button press response. They received two blocks of 18 trials each. Arrow direction was counterbalanced across each block of 18 trials and presented in random order. The computer signaled incorrect responses and recorded latency (in milliseconds) and accuracy of responses. All children met the criterion of at least 80% correct responses within the 36 response mapping trials.

Recentration (updating of working memory)

Recentration (updating) was assessed with the *adapted letter memory task* (Agostino et al., 2010), which includes 4 practice trials and 16 test trials. Within each trial, individual letters were presented serially on a computer screen at the rate of 2000 ms per letter. The list length varied randomly over the trials and contained 5, 7, 9, or 11 letters. To ensure that children were updating continuously, they were instructed to orally recall the last 3 letters presented on the monitor. This required children to add the most recent letter presented while dropping the fourth letter back as they worked through each list. The score was the proportion of correct answers on the test trial items that required updating ($n = 80$).

Decentration (shifting of mental sets)

Decentration (shifting) was measured using the *contingency naming task* (CNT; Anderson, Anderson, Northam, & Taylor, 2000). The contingency naming task includes four subtests: (a) baseline color naming task, (b) baseline shape naming task, (c) a one-dimensional switching task, and (d) a two-dimensional switching task. The stimulus consists of a card with three rows of shapes (circle, square, or triangle) of different colors (pink, blue, or green). Within each outer shape, a second inner shape is drawn. Above nine of the shapes is a reverse arrow. For the first subtest children were asked to name the color of the shape, and for the second subtest children were asked to name the outer shape. Subtests 3 and 4 involved a shifting of mental sets from shape to color or from color to shape. For subtest 3, children were provided with the following rule: Name the color if the inner and outer shapes match or the outer shape if they do not. For subtest 4, children were instructed to apply the same rule used in subtest 3 unless the shape had a backward arrow over it. Where a backward arrow was present, children were required to reverse the rule from subtest 3 (i.e., to name the color if the inner and outer shapes were different). This task has been used to assess switching between response sets in children as young as 6 years (Mazzocco & Kover, 2007). For all trials, children were asked to complete the task as quickly as possible using their finger to point to each shape. Completion time (in seconds) and errors were recorded for each subtest. An efficiency score that reflects both speed and accuracy was used to index shifting ability. The following formula was used to calculate the efficiency score: $[(1/\text{total completion time})/\sqrt{(\text{total errors} + 1)}] \times 100$, where higher scores represented better shifting ability (Anderson et al., 2000).

Theory of mind

The Happé Strange Stories task (Happé, 1994) was used to assess *higher order ToM*. The Strange Stories task is a widely used measure of higher order ToM that is composed of 12 short stories with varying, but realistic, content that require understanding of others' intentions and motives (Devine & Hughes, 2013; Gillott et al., 2004; Happé, 1994; Liddle & Nettle, 2006). Based on pilot research (Im-Bolter, Owens, & Bauer, 2009), 5 of the 12 stories were excluded from the current study because they were determined to be the least reliable due to too much ambiguity in the stories and/or content that could be interpreted in several different ways (e.g., child wearing a ghost costume for Halloween, child forgetting that his classmate was at school that day). The 7 stories included represented the following story types: Joke, Persuasion, Figure of Speech, Lie, White Lie, Contrary Emotions, and

Misunderstanding. The stories were read orally and presented in print in order to reduce memory demands. Each story was accompanied by two questions. The first question assessed story comprehension—"Was it true what X said?"—and was scored as correct or incorrect. The second question assessed whether children were able to determine the protagonist's mental intention—"Why did X say that?"—and was used to assess higher order ToM. The answers to the ToM question were scored on a scale ranging from 0 to 2.5, where 0 represented an incorrect response, 1 represented a response that referred to physical or literal states (e.g., "Her hair is too short," "She looks silly"), 2 represented an answer that referred to the mental state (e.g., thoughts, feelings, traits, dispositions) of the protagonist and included mental state terms (e.g., "He *wanted* to be funny"), and a 2.5 represented a response that referred to the mental state of both characters in the story (e.g., "He *thought* it was funny and he was making a joke to his friend, who probably *thought* it was funny too"). This resulted in a total possible score of 17.5 across the 7 stories. The ToM score was calculated as proportion correct.

Interrater reliabilities of at least .80 were confirmed on a sample of data not included in the current study before scoring of the study data commenced. The kappa coefficients (all $ps < .0001$) for the 7 stories (in the order listed above) are as follows: $\kappa = .89$, $\kappa = 1.00$, $\kappa = .96$, $\kappa = .91$, $\kappa = .91$, $\kappa = .87$, and $\kappa = 1.00$. At fixed intervals, reliability checks were conducted on approximately 20% of the stories to ensure continued reliability. Disagreements were resolved through discussion.

Procedure

Participants received four individual sessions. The first session comprised the PIQ and language screening measures. All children received the intelligence screening first to ensure that they met PIQ criteria. The remaining sessions were given in a random order: (a) Happé Strange Stories and adapted letter memory task, (b) antisaccade task and contingency naming task, and (c) figural intersections test and direction following task. Children were tested in their schools during school hours and were provided with breaks as necessary. Children did not receive more than one session per day. At the end of each session, children were provided with a token of appreciation for their participation (e.g., stickers).

Results

Data screening

Before analyses, we examined data distributions for normality, sphericity, skewness, and kurtosis. All measures met criteria for multivariate normality (Kline, 2005). For path analyses, we examined bivariate, residual, and influence plots. All variables appeared to have linear relationships. Multivariate outliers did not seem to have an undue influence on the variables with the exception of 2 children in the middle childhood group. Examination of notes indicated that 1 child was experiencing hearing problems and the other child was sleepy during testing. As a result, both children were excluded from the analyses and the final sample for the middle childhood group included a total of 226 children aged 7 or 8 years (119 boys and 107 girls).

Sample characteristics

Analysis of variance (ANOVA) was used to assess group differences on measures of SES, estimated PIQ, and language. Chi-square analysis was used to assess group differences in gender. As can be seen in Table 1, the groups did not differ significantly with respect to gender, SES, or syntactic language skills. The middle childhood group showed significantly higher estimated PIQ and semantic language skills; however, the effect size for these variables was very small (η^2 values = .014 and .043, respectively). Regardless, an analysis of covariance (ANCOVA) with PIQ and semantic language skills as covariates was used to assess group differences for the variables below, and adjusted means are reported.

Table 1
Group differences for SES, estimated PIQ, language, and *M* capacity.

	Childhood ($n = 226$)		Adolescence ($n = 216$)		F/χ^2 ($df = 1, 440$)	p	η^2
	M	*SD*	*M*	*SD*			
Boys *n* (%)	119 (52.65)		106 (49.10)		0.57	.452	
Age (years)	8.08	0.51	12.09	0.49			
(range)	(7.00–8.92)		(11.00–12.92)				
SES[a]	50.79	11.76	50.31	13.47	0.15	.696	.000
Estimated PIQ	104.44	13.26	101.54	10.47	6.45	.011	.014
Language							
Semantics							
PPVT-4	111.49	11.45	105.03	10.90	36.82	.0001	.077
EVT-2	106.18	11.31	104.05	10.84	4.08	.044	.009
Syntax							
TROG-2	97.34	12.95	98.32	10.12	0.79	.376	.002
CELF-4-FS	11.22	2.17	11.00	2.34	1.36	.303	.002
Overall semantic	.59	.70	.30	.66	19.73	.0001	.043
Overall syntactic	.12	.68	.11	.60	0.00	.947	.000
M capacity					($df = 3, 438$)		
FIT *M* score	2.82	1.06	4.44	1.34	214.95	.0001	.314
DFT *M* score	3.05	1.10	4.44	0.93	222.18	.0001	.323
Composite *M* score	2.93	0.87	4.44	0.89	376.92	.0001	.437

Note: SES, socioeconomic status; PPVT-4, Peabody Picture Vocabulary Test–Fourth Edition; EVT-2, Expressive Vocabulary Test–Second Edition; TROG-2, Test for the Reception of Grammar–Second Edition; CELF-4-FS, Clinical Evaluation of Language Fundamental–Fourth Edition–Formulated Sentences; FIT, figural intersections test; DFT, direction following task. PIQ and language scores are standard scores, with the exception of overall semantic and overall syntactic language scores, which are *z* scores. Effect size = η^2, proportion of total variance accounted for by the group effect.

[a] $n = 217$ for childhood group and $n = 207$ for early adolescence group.

Mental attentional capacity (M capacity)

Mean scores for the FIT, DFT, and composite *M* score appear in Table 1. The age of the middle childhood group (7 and 8 years) corresponds to a theoretically expected *M* score of 3, and the age of the early adolescence group (11 and 12 years) corresponds to a theoretically expected *M* score of 5. The *M* scores show that both groups are close to what is theoretically expected for their age, although the early adolescence group demonstrates slight underperformance on measures of *M* capacity.

Attentional interruption (inhibition)

As expected, analysis of mean accuracy on the antisaccade task showed that the middle childhood group ($M = 46.66$, $SD = 14.25$) demonstrated worse inhibition skills than the early adolescence group ($M = 73.87$, $SD = 10.47$), $F(3, 438) = 498.62$, $p < .0001$, $\eta^2 = .528$.

Executive processes

The middle childhood group had significantly lower accuracy ($M = 18.10$, $SD = 14.83$) on the updating trials of the adapted letter memory task compared with the early adolescence group ($M = 42.01$, $SD = 17.31$), $F(3, 438) = 401.64$, $p < .0001$, $\eta^2 = .473$. The middle childhood group also was less efficient on the contingency naming task ($M = .24$, $SD = .13$) compared with the early adolescence group ($M = .44$, $SD = .16$), $F(3, 438) = 232.31$, $p < .0001$, $\eta^2 = .340$. These results indicate that the middle childhood group shows more difficulty with both recentration (updating) and decentration (shifting) compared with the early adolescence group.

Theory of mind

The middle childhood group had a significantly lower score ($M = 60.84$, $SD = 11.22$) on the Happé Strange Stories task compared with the early adolescence group ($M = 70.89$, $SD = 6.76$), $F(1, 440) = 137.68$, $p < .0001$, $\eta^2 = .229$.

Predicting higher order ToM in middle childhood

Path analysis was used to test the theoretical model illustrated in the top panel of Fig. 2. We used the following variables: (a) *M* capacity, which was indexed by a composite *M* score (mean *M* score from the FIT and DFT), $r(226) = .30$, $p < .0001$; (b) higher order ToM, which was represented by the proportion correct score from the Happé Strange Stories; (c) *I* (inhibition), which was indexed by accuracy on the antisaccade task; (d) decentration (shifting), which was reflected by the efficiency score on the contingency naming task; (e) recentration (updating), which was represented by the proportion of correct updating trials on the adapted letter memory task; (f) syntactic language, which was reflected by a standardized composite score (mean *z* score based on the normative mean) from the two syntactic language tests, $r(226) = .46$, $p < .0001$; and (g) semantic language, which was represented by a standardized composite score (mean *z* score based on the normative mean) from the two semantic tests, $r(226) = .69$, $p < .0001$.

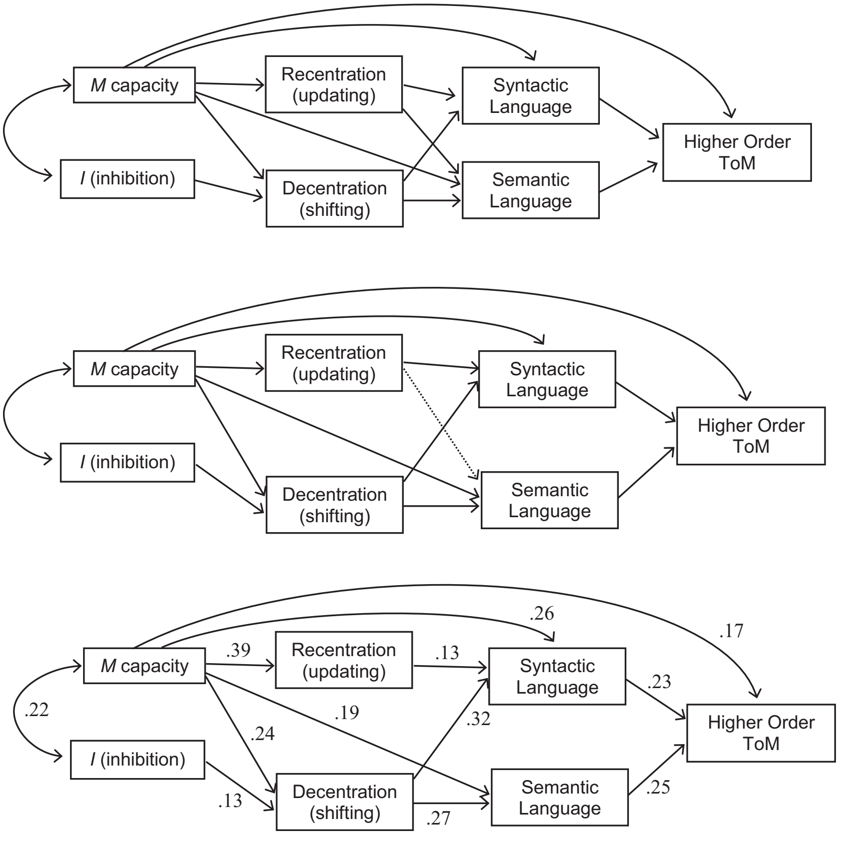

Fig. 2. Middle childhood model. Top panel: Initial theoretical model. Middle panel: Theoretical model with significant paths indicated by solid lines. Bottom panel: Revised theoretical model with standardized path coefficients; all paths significant.

Means, standard deviations, and intercorrelations of the relevant variables for the middle childhood group are shown above the diagonal in Table 2. Note that *M* capacity ($r = .29$), higher order ToM ($r = .40$), *I* (inhibition) ($r = .19$), and recentration ($r = .26$) were all significantly correlated with age ($p = .005$ or $p < .0001$). The correlation between decentration and age approached significance ($r = .12$, $p = .066$). We used the SAS CALIS procedure to conduct a path analysis. The initial model yielded a good fit to the data, as indicated by a nonsignificant $\chi^2(6, N = 226) = 11.75$, $p = .065$, small residuals (root mean square error of approximation or RMSEA = .065), and large fit index values in excess of .95 (normed fit index or NFI = .96, comparative fix index or CFI = .98, and goodness-of-fit index or GFI = .99). All path coefficients were significant at the .05 level with the exception of the path between recentration (updating) and semantic language (see middle panel of Fig. 2). A revised model removing this path was examined and yielded a good fit to the data, as indicated by a nonsignificant $\chi^2(7, N = 226) = 11.86$, $p = .105$, small residuals (RMSEA = .056), and large fit index values in excess of .95 (NFI = .96, CFI = .98, GFI = .98). All path coefficients were significant at the .05 level. The revised model appears in the bottom panel of Fig. 2.

By comparing the chi-square statistics for the initial model with that for the revised model, it was possible to perform a chi-square difference test to determine whether removal of the insignificant path resulted in a significant decrease in model fit. This difference was computed as 11.86 − 11.75 = 0.11. With $df = 1$, the chi-square difference statistic was not significant ($p = .740$), indicating that the revised model provided an equally good fit to the data, with the advantage of being the more parsimonious model.

Predicting higher order ToM in early adolescence

Path analysis was used to test the theoretical model illustrated in the top panel of Fig. 3. The same variables, with the exception of recentration (shifting), were used as described above for the middle childhood group. Means, standard deviations, and intercorrelations of the relevant variables for the early adolescence group are shown below the diagonal in Table 2. Note that *M* capacity, higher order ToM, decentration, and recentration were not significantly correlated with age; however, the correlation between age and *I* (inhibition) approached significance ($r = .13$, $p = .051$). We used the SAS CALIS procedure to conduct a path analysis. The initial model offered a good fit to the data, as indicated by a nonsignificant $\chi^2(6, N = 216) = 3.64$, $p = .725$, small residuals (RMSEA = .000), and large fit index values in excess of .95 (NFI = .97, CFI = 1.00, GFI = .99). All path coefficients were significant at the .05 level with the exception of the paths between syntactic language and higher order ToM and between decentration (shifting) and syntactic language (see middle panel of Fig. 3). A revised model removing

Table 2
Means, standard deviations, and intercorrelations for measures of higher order ToM, *M* capacity, *I* (inhibition), decentration (shifting), recentration (updating), syntactic language, and semantic language as a function of group.

Measure	1	2	3	4	5	6	7	*M*	*SD*
1. ToM	–	.32***	.16+	.23**	.30***	.41***	.41***	.61	.11
2. *M* capacity	.08	–	.22***	.27**	.39***	.40***	.26***	3.01	0.87
3. *I* (inhibition)	.04	.24**	–	.22**	.21*	.12	.07	.47	.14
4. Decentration (shifting)	.22**	.21*	.07	–	.34***	.44***	.32***	.25	.12
5. Recentration (updating)	–	–	–	–	–	.33***	.15+	.24	.15
6. Syntactic	.18*	.27***	.02	.20*	–	–	.49***	.12	.68
7. Semantic	.21*	.34***	.12	.15+	–	.49***	–	.59	.70
M	.70	4.36	.73	.44	–	.11	.30		
SD	.68	.89	.10	.16	–	.60	.66		

Note: Intercorrelations for the middle childhood group ($n = 226$) are above the diagonal, and intercorrelations for the early adolescence group ($n = 216$) are below the diagonal. Means and standard deviations for the middle childhood group are in the vertical columns, and means and standard deviations for the early adolescence group are in the horizontal rows.

+ $p < .05$.
* $p < .01$.
** $p < .001$.
*** $p < .0001$.

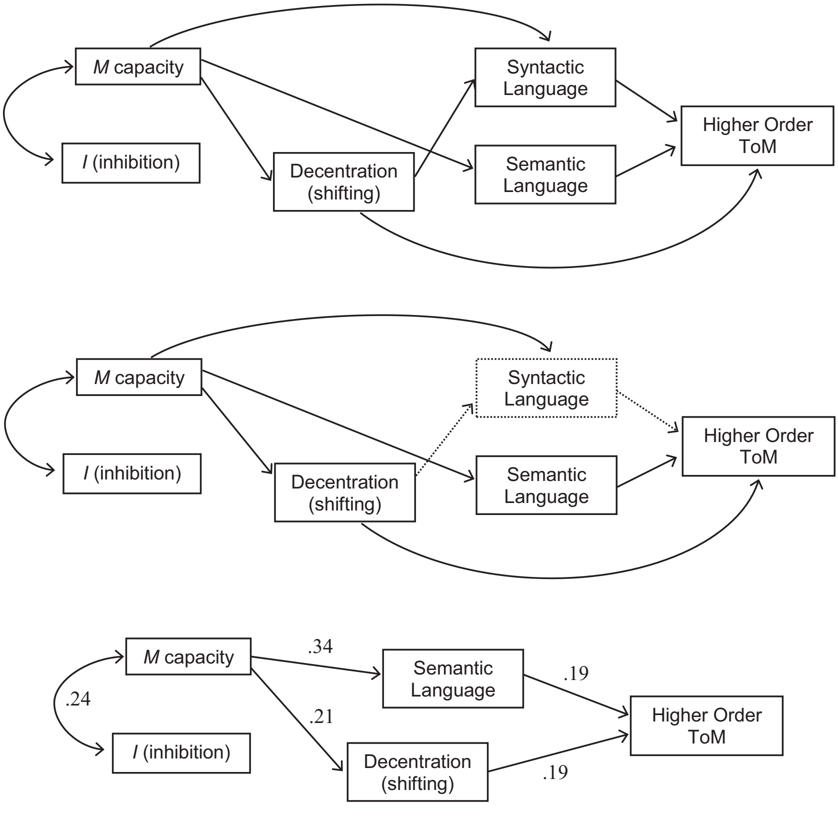

Fig. 3. Early adolescence model. Top panel: Theoretical model. Middle panel: Theoretical model with significant paths indicated by solid lines. Bottom panel: Revised theoretical model with standardized path coefficients; all paths significant.

syntactic language was examined and yielded a good fit to the data, as indicated by a nonsignificant $\chi^2(5, N = 216) = 2.27$, $p = .810$, small residuals (RMSEA = .000), and large fit index values in excess of .95 (NFI = .97, CFI = 1.00, GFI = 1.00). All path coefficients were significant at the .05 level. The revised model appears in the bottom panel of Fig. 3.

By comparing the chi-square statistics for the initial model with that for the revised model, it was possible to perform a chi-square difference test to determine whether removal of the insignificant paths resulted in a significant decrease in model fit. This difference was computed as 3.61 − 2.27 = 1.37. With $df = 1$, the chi-square difference statistic was not significant ($p = .242$), indicating that the revised model provided an equally good fit to the data, with the advantage of being the more parsimonious model.

Discussion

Both language and EF are important correlates of first-order ToM in the preschool period; however, little is known about the continuity of these relations or their pattern of associations as children grow older. The primary goal of the current study was to use the theory of constructive operators to test two models of higher order ToM: one for middle childhood and the other for early adolescence. The results of our study show that differences in higher order ToM between middle childhood and early adolescence are associated with gains in *M* capacity, *I* (inhibition), EF, and language. Moreover, our findings indicate that the model of associations in these two age groups reflect involvement of different cognitive resources. We propose that this is related to the different mental demands experienced by, and level of mental resources available to, children and adolescents during processing of mental perspectives. Results from the current study show increases in *M* capacity, *I* (inhibition), EF, language, and higher order ToM from middle childhood to early adolescence. These increases are consistent with the theory of constructive operators, our understanding of child and adolescent development, and previous research (e.g., Best & Miller, 2010; Devine & Hughes, 2013; Huizinga, Dolan, & van der Molen, 2006; Im-Bolter et al., 2006; Johnson et al., 2003).

Models of ToM: Middle childhood

The results of the path analysis support a model of higher order ToM in middle childhood that reflects more intensive involvement of cognitive resources (particularly *M* capacity) compared with the model for early adolescence. This is likely due to the relative novice level of 7- and 8-year-olds with respect to higher order ToM. Our findings also provide an integrated extension of what we know about first-order ToM in younger children. Both language and EF continue to be important for higher order ToM and have an interactive relation that contributes to successful mental perspective taking. Our results suggest that one is not more important than the other for higher order ToM but rather that they work together. This may help to explain why Chinese preschoolers show advanced EF but not advanced ToM when compared with their U.S. counterparts (Sabbagh, Xu, Carlson, Moses, & Lee, 2006). In addition, syntactic and semantic language appear to have different associations with EF in middle childhood. Although decentration (shifting) is related to syntactic and semantic language, recentration (updating) has an association with syntactic language only. This differential relation may be the result of a trade-off between increased language competence and decreased language demand in mental perspective taking in this age range. Although updating of sentence structures might assist in considering different mental perspectives, there is no need to update word meaning because activation of all relevant schemes is maintained (either by *M* or by the situation). Shifting, however, is a crucial aspect of coordinating different perspectives and, therefore, is associated with both aspects of language, semantic and syntactic, during the process of considering mental intentions.

Our results indicate that *M* capacity is an important component in our model of higher order ToM in middle childhood. It has both a direct and indirect path (via language and EF) to higher order ToM, which suggests that *M* capacity plays a primary role in optimizing the ability to consider mental intentions when predicting or explaining behavior in middle childhood. This is consistent with the proposal that increases in processing capacity underlie advancements in ToM (Gordon & Olson, 1998) and makes sense when we consider that the mental demand of higher order ToM tasks may be higher than, or at the same level as, a child's *M* capacity. However, as children get older and their *M* capacity "catches up" to the mental demand of perspective taking, a different model of associations may emerge for ToM.

Models of ToM: Early adolescence

Results of the current study indicate a less complex model of associations for higher order ToM in early adolescence. As children get older, they have greater cognitive resources (e.g., *M* capacity) available to them and more experience with ToM. This developmental change may result in a different pattern of processes that predict ToM because the mental demand of perspective taking is within their *M* capacity. In early adolescence, *M* capacity appears to have no role in higher order ToM (i.e., no zero-order correlation); however, *M* capacity has a significant relation to both semantic language and decentration (shifting), both of which are important for higher order ToM in this age group. The results of the path analysis in early adolescence indicate that the influence of *M* capacity is subtle compared with that seen in middle childhood and not noticeable unless examined simultaneously with semantic language and decentration (shifting). Our findings suggest that decentration (shifting) has a more prominent role in higher order ToM as children get older, which is consistent with the proposal that ToM may increasingly depend on executive processes in adults (Apperly et al., 2009). The ability to maintain activation of both perspectives (via *M* capacity) while shifting between them allows an individual to directly compare and contrast different viewpoints in order to determine and evaluate the intentions behind behavior. Anecdotally, this is something we increasingly see in adolescence.

An interesting aspect of our early adolescence model of higher order ToM is the continued importance of semantic language for perspective taking. It is possible that syntactic language competence is important during the emergent stage of higher order ToM because it requires syntactic structures that can represent modified, elaborated, or contrastive mental states (often signaled by conjunctions). This may parallel the importance of sentential complements for first-order ToM (false beliefs) in the preschool period. Once the individual moves beyond the novice period, syntactic language demands may plateau and representation of mental perspectives may no longer depend on syntactic competence. Neuropsychological research suggesting that ToM may be independent of grammar in adults (see Apperly et al., 2009, for a review) is consistent with this idea. Semantic language, however, continues to grow in a complex, and sometimes subtle, manner well into adulthood (e.g., inferential or nonliteral meaning, complex emotions, conditional meanings). Semantic language may provide the vehicle for representation of increasingly nuanced and sophisticated intentions, beliefs, and emotions. The model of higher order ToM in early adolescence examined in the current study indicates that this may be a time of transition toward a mature ToM system seen in adults. The scores achieved on the higher order ToM task by the early adolescence group are not anywhere near ceiling levels, suggesting that significant improvement can still be made. In addition, there are clearly other factors that contribute to higher order ToM that we did not include (e.g., other aspects of language such as pragmatic, figurative, and social discourse; social environmental factors such as number of siblings and reading experiences) that may be critical in early adolescence but perhaps less so in middle childhood. For example, the capacity to understand and produce figurative language increases rapidly in adolescence and becomes a prominent feature in exchanges with peers. As a result, it is possible that figurative language is an important component in a model of higher order ToM in adolescence compared with one in middle childhood.

We acknowledge that we need to be cautious about drawing any causal conclusions because the data were collected contemporaneously. Research using data from longitudinal and intervention studies is needed to clarify the causal relations that exist among the different variables. In addition, the models tested in the current study need to be replicated with children from different backgrounds, cultures, and development (e.g., children with specific language impairment or with autism). This may help to explain how differences in specific processes are associated with performance on ToM tasks.

Conclusions

The current study represents the first to test models of higher order ToM that examine *M* capacity, *I* (inhibition), language, and EF simultaneously in middle childhood and early adolescence in a theory-guided manner. Our results suggest that different systems of processes are involved in the two age groups and demonstrate the importance of considering how mental demand and mental processes interact in higher order ToM at different stages of development. In middle childhood a complex system of cognitive resources, which includes *M* capacity, language, and EF, appears to contribute to higher order ToM, whereas in early adolescence the primary contributors seem to be semantic language and decentration (shifting of mental sets). These findings highlight the importance of integrating different cognitive and linguistic processes into a single model in order to understand their contribution to ToM at different developmental stages.

Acknowledgments

An aspect of this article was based on a master's thesis submitted by Keely Owens-Jaffray to the Applied Modeling and Quantitative Methods Graduate Program at Trent University. Portions of this research were presented at the 2011 meeting of the Society for Research in Child Development in Montreal, Quebec, Canada, and the 2012 meeting of the Jean Piaget Society in Toronto, Ontario, Canada. The research was supported by a Social Sciences and Humanities Research Council of Canada

(SSHRC) standard research grant (410-2009-0776) to Nancie Im-Bolter. We thank the children and parents who participated in this project.

References

Agostino, A., Johnson, J., & Pascual-Leone, J. (2010). Executive functions underlying multiplicative reasoning: Problem type matters. *Journal of Experimental Child Psychology, 105*, 286–305.
Ahmed, F. S., & Miller, S. (2011). Executive function mechanisms in theory of mind. *Journal of Autism and Developmental Disorders, 41*, 667–678.
Anderson, P., Anderson, V., Northam, E., & Taylor, H. (2000). Standardization of the contingency naming test for school-aged children: A new measure of reactive flexibility. *Clinical Neuropsychological Assessment, 1*, 247–273.
Apperly, I. A., Carroll, D. J., Samson, D., Humphreys, G. W., Qureshi, A., & Moffitt, G. (2010). Why are there limits on theory of mind use? Evidence from adults' ability to follow instructions from an ignorant speaker. *Quarterly Journal of Experimental Psychology, 63*, 1201–1217.
Apperly, I. A., Samson, D., & Humphreys, G. W. (2009). Studies of adults can inform accounts of theory of mind development. *Developmental Psychology, 45*, 190–201.
Apperly, I. A., Warren, F., Andrews, B. J., Grant, J., & Todd, S. (2011). Developmental continuity in theory of mind: Speed and accuracy of belief–desire reasoning in children and adults. *Child Development, 82*(5), 1691–1703. http://dx.doi.org/10.1111/j.1467-8624.2011.01635.x.
Arsalidou, M., Pascual-Leone, J., & Johnson, J. (2010). Misleading cues improve developmental assessment of working memory capacity: The color matching tasks. *Cognitive Development, 25*, 262–277.
Astington, J. W., & Jenkins, J. M. (1999). A longitudinal study of the relation between language and theory of mind development. *Developmental Psychology, 35*, 1311–1320.
Austin, G., Gropper, 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*. http://dx.doi.org/10.3389/fpsyg.2015.000655.
Baddeley, A. (1996). Exploring the central executive. *Quarterly Journal of Experimental Psychology A: Human Experimental Psychology, 49*, 5–28.
Benson, J. E., Sabbagh, M. A., Carlson, S. M., & Zelazo, P. D. (2013). Individual differences in executive functioning predict preschoolers' improvement from theory-of-mind training. *Developmental Psychology, 49*, 1615–1627.
Best, J. R., & Miller, P. H. (2010). A developmental perspective on executive function. *Child Development, 81*, 1641–1660.
Bishop, D. V. M. (2003). *Test for reception of grammar-second edition*. Manchester, UK: University of Manchester.
Blishen, B., Carroll, W., & Moore, C. (1987). The 1981 socioeconomic index for occupations in Canada. *Canadian Review of Sociology, 24*, 465–488.
Bosacki, S. L. (2000). Theory of mind and self-concept in preadolescents: Links with gender and language. *Journal of Educational Psychology, 92*, 709–717.
Brydges, C. R., Reid, C. L., Fox, A. M., & Anderson, M. (2012). A unitary executive function predicts intelligence in children. *Intelligence, 40*, 458–469.
Carlson, S. M., Mandell, D. J., & Williams, L. (2004). Executive function and theory of mind: Stability and prediction from ages 2 to 3. *Developmental Psychology, 40*, 1105–1122.
Carlson, S. M., & Moses, L. J. (2001). Individual differences in inhibitory control and children's theory of mind. *Child Development, 72*, 1032–1053.
Carlson, S. M., Moses, L. J., & Breton, C. (2002). How specific is the relation between executive function and theory of mind? Contributions of inhibitory control and working memory. *Infant and Child Development, 11*, 73–92.
Cunning, S. (2003). *The direction following task: Assessing mental capacity in the linguistic domain*. Unpublished doctoral dissertation, York University, Toronto, Ontario, Canada.
de Villiers, J. G., & Pyers, J. E. (2002). Complements to cognition: A longitudinal study of the relationship between complex syntax and false-belief understanding. *Cognitive Development, 17*, 1037–1060.
Devine, R. T., & Hughes, C. (2013). Silent films and strange stories: Theory of mind, gender, and social experiences in middle childhood. *Child Development, 84*, 989–1003.
Devine, R. T., & Hughes, C. (2014). Relations between false belief understanding and executive function in early childhood: A meta-analysis. *Child Development, 85*, 1777–1794.
Dumontheil, I., Apperly, I. A., & Blakemore, S. (2010). Online usage of theory of mind continues to develop in late adolescence. *Developmental Science, 13*, 331–338.
Dunn, L. M., & Dunn, L. M. (2006). *The Peabody picture vocabulary test-fourth edition*. Circle Pines, MN: American Guidance Service.
Engle, R. W., Conway, A. R., Tuholski, S. W., & Shisler, R. J. (1995). A resource account of inhibition. *Psychological Science, 6*, 122–125.
Farrant, B. M., Fletcher, J., & Maybery, M. T. (2006). Specific language impairment, theory of mind, and visual perspective taking: Evidence for simulation theory and the developmental role of language. *Child Development, 77*, 1842–1853.
Farrant, B. M., Maybery, M. T., & Fletcher, J. (2012). Language, cognitive flexibility, and explicit false belief understanding: Longitudinal analysis in typical development and specific language impairment. *Child Development, 83*, 223–235.
Gillott, A., Furniss, F., & Walter, A. (2004). Theory of mind ability in children with specific language impairment. *Child Language Teaching and Therapy, 20*(1), 1–11.
Gordon, A., & Olson, D. (1998). The relation between acquisition of a theory of mind and the capacity to hold in mind. *Journal of Experimental Child Psychology, 68*, 70–83.
Happé, F. (1994). An advanced test of theory of mind: Understanding of story characters' thoughts and feelings by able autistic, mentally handicapped, and normal children and adults. *Journal of Autism and Developmental Disorders, 24*, 129–154.
Howard, S. J., Johnson, J., & Pascual-Leone, J. (2014). Clarifying inhibitory control: Diversity and development of attentional inhibition. *Cognitive Development, 31*, 1–21. http://dx.doi.org/10.1016/j.cogdev.2014.03.001.

Hughes, C. (1998). Executive function in preschoolers: Links with theory of mind and verbal ability. *British Journal of Developmental Psychology, 16*, 233–253.
Huizinga, M., Dolan, C. V., & van der Molen, M. W. (2006). Age-related change in executive function: Developmental trends and a latent variable analysis. *Neuropsychologia, 44*, 2017–2036.
Im-Bolter, N., Owens, K., & Bauer, S. (2009, April). *Theory of mind in children with language impairment*. Poster presented at biennial meeting of the Society for Research in Child Development, Denver CO.
Im-Bolter, N., Johnson, J., Ling, D., & Pascual-Leone, J. (2015). Inhibition: Mental control process or mental resource? *Journal of Cognition and Development, 16*, 666–681.
Im-Bolter, N., Johnson, J., & Pascual-Leone, J. (2006). Processing limitations in children with specific language impairment: The role of executive function. *Child Development, 77*, 1822–1841.
Johnson, J., Im-Bolter, N., & Pascual-Leone (2003). Development of mental attention in gifted and mainstream children: The role of mental capacity, inhibition, and speed of processing. *Child Development, 74*, 1594–1614.
Keysar, B., Lin, S., & Barr, D. J. (2003). Limits on theory of mind use in adults. *Cognition, 89*, 25–41.
Kline, R. B. (2005). *Principles and practice of structural equation modeling* (2nd Ed.). New York: Guilford Press.
Lehto, J. E., Juujarvi, P., Kooistra, L., & Pulkkinen, L. (2003). Dimensions of executive functioning: Evidence from children. *British Journal of Developmental Psychology, 21*, 59–80.
Liddle, B., & Nettle, D. (2006). Higher-order theory of mind and social competence in school-age children. *Journal of Cultural and Evolutionary Psychology, 4*, 231–246.
Low, J. (2010). Preschoolers' implicit and explicit false-belief understanding: Relations with complex syntactical mastery. *Child Development, 81*, 597–615.
Markel, C., Major, A., & Pelletier, J. (2013). Relations among children's theory of mind, family factors, language development, and story comprehension in L1 and L2 preschoolers. *International Journal of Research Studies in Psychology, 2*(2), 15–32.
Mazzocco, M. M., & Kover, S. T. (2007). A longitudinal assessment of executive function skills and their association with math performance. *Child Neuropsychology, 13*, 18–45.
Miller, C. (2001). False belief understanding in children with specific language impairment. *Journal of Communication Disorders, 34*, 73–86.
Miller, S. (2009). Children's understanding of second-order mental states. *Psychological Bulletin, 135*, 749–773.
Miller, M. R., Giesbrecht, G. F., Müller, U., McInerney, R. J., & Kearns, K. A. (2012). A latent variable approach to determining the structure of executive function in preschool children. *Journal of Cognition and Development, 13*, 395–423.
Miller, R., Pascual-Leone, J., Campbell, C., & Juckes, T. (1989). Cross-cultural similarities and differences on two neo-Piagetian cognitive tasks. *International Journal of Psychology, 24*, 293–313.
Milligan, K., Astington, J. W., & Dack, L. A. (2007). Language and theory of mind: Meta-analysis of the relation between language ability and false-belief understanding. *Child Development, 78*, 622–646.
Miyake, A., & Friedman, N. P. (2012). The nature and organization of individual differences in executive functions: Four general conclusions. *Current Directions in Psychological Science, 21*, 8–14.
Miyake, A., Friedman, N. P., Emerson, M. J., Witzki, A. H., Howerter, A., & Wager, T. D. (2000). The unity and diversity of executive functions and their contributions to complex "frontal lobe" tasks: A latent variable analysis. *Cognitive Psychology, 41*, 49–100.
Müller, U., Liebermann-Finestone, D. P., Carpendale, J. I. M., Hammond, S. I., & Bibok, M. B. (2012). Knowing minds, controlling actions: The developmental relations between theory of mind and executive function from 2 to 4 years of age. *Journal of Experimental Child Psychology, 111*, 331–348.
Pascual-Leone, J. (1970). A mathematical model for the transition rule in Piaget's developmental stages. *Acta Psychologica, 32*, 301–345.
Pascual-Leone, J. (1987). Organismic processes for neo-Piagetian theories: A dialectical causal account of cognitive development. *International Journal of Psychology, 22*, 531–570.
Pascual-Leone, J., & Ijaz, H. (1989). Mental capacity testing as a form of intellectual-developmental assessment. In R. Samuda, S. Kong, J. Cummings, J. Pascual-Leone, & J. Lewis (Eds.), *Assessment and placement of minority students* (pp. 144–171). Lewiston, NY: Hogrefe.
Pascual-Leone, J., & Johnson, J. (2001). *Manual for the FIT: Figural intersections test*. Toronto, Ontario, Canada: Developmental Processes Laboratory, Department of Psychology, York University.
Pascual-Leone, J., & Johnson, J. (2011). A developmental theory of mental attention: Its application to measurement and task analysis. In P. Barrouillet & V. Gaillard (Eds.), *Cognitive development and working memory: A dialogue between neo-Piagetian theories and cognitive approaches* (pp. 13–46). New York: Psychology Press.
Pascual-Leone, J., & Johnson, J. (1999). A dialectical constructivist view of representation: Role of mental attention, executives, and symbols. In S. Irving (Ed.), *Development of mental representation: Theories and applications* (pp. 169–200). Mahwah, NJ: Lawrence Erlbaum.
Pascual-Leone, J., & Johnson, J. (2005). A dialectical constructivist view of developmental intelligence. In O. Wilhelm & R. W. Engle (Eds.), *Handbook of understanding and measuring intelligence* (pp. 177–201). Thousand Oaks, CA: Sage.
Perner, J., & Lang, B. (1999). Development of theory of mind and executive control. *Trends in Cognitive Sciences, 3*, 337–344.
Powell, T. L., Arsalidou, M., Vogan, V., & Taylor, M. J. (2014). Letter and color matching tasks: Parametric measure of developmental working memory capacity. *Child Development Research, 2014*. http://dx.doi.org/10.1155/2014/961781.
Roberts, R. J., Hager, L. D., & Heron, C. (1994). Prefrontal cognitive processes: Working memory and inhibition in the antisaccade task. *Journal of Experimental Psychology: General, 123*, 374–393.
Rose, S., Feldman, J. F., & Jankowski, J. J. (2011). The structure of memory in infants and toddlers: An SEM student with full-terms and preterms. *Developmental Science, 14*, 83–91.
Ruffman, T., Slade, L., Rowlandson, K., Rumsey, C., & Garnham, A. (2003). How language relates to belief, desire, and emotion understanding. *Cognitive Development, 18*, 139–158.
Sabbagh, M. A., Xu, F., Carlson, S., Moses, L. J., & Lee, K. (2006). The development of executive functioning and theory of mind: A comparison of Chinese and U.S. preschoolers. *Psychological Science, 17*, 74–81.

Semel, E., Wiig, E., & Secord, W. (2003). *Clinical evaluation of language fundamentals-fourth edition*. San Antonio, TX: Harcourt Brace.
Tager-Flusberg, H., & Joseph, R. M. (2005). How language facilitates the acquisition of false belief understanding in children with autism. In J. Astington & J. Baird (Eds.), *Why language matters for theory of mind* (pp. 298–318). Oxford, UK: Oxford University Press.
Vetter, N. C., Altgassen, M., Phillips, L., Mahy, C. E. V., & Kliegel, M. (2013). Development of affective theory of mind across adolescence: Disentangling the role of executive functions. *Developmental Neuropsychology, 38*, 114–125.
Vetter, N. C., Leipold, K., Kliegel, M., Phillips, L. H., & Altgassen, M. (2013). Ongoing development of social cognition in adolescence. *Child Neuropsychology, 19*, 615–629.
Wechsler, D. (1999). *Wechsler abbreviated scale of intelligence*. San Antonio, TX: Harcourt Brace.
Williams, K. T. (2006). *Expressive vocabulary test-second edition*. Circle Pines, MN: American Guidance Services.