Neural Networks 23 (2010) 1043–1050

Contents lists available at ScienceDirect

Neural Networks

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

2010 Special Issue

Conceptual change and preschoolers' theory of mind: Evidence from load–force adaptation

Mark A. Sabbagh*, Sydney F.R. Hopkins, Jeannette E. Benson, J. Randall Flanagan

Psychology Department, Queen's University, Kingston K7L 3N6, ON, Canada

ARTICLE INFO

Keywords:
Theory of mind
Load–force adaptation
Theory theory
Social cognitive development
Developmental cognitive neuroscience

ABSTRACT

Prominent theories of preschoolers' theory of mind development have included a central role for changing or adapting existing conceptual structures in response to experiences. Because of the relatively protracted timetable of theory of mind development, it has been difficult to test this assumption about the role of adaptation directly. To gain evidence that cognitive adaptation is particularly important for theory of mind development, we sought to determine whether individual differences in cognitive adaptation in a non-social domain predicted preschoolers' theory of mind development. Twenty-five preschoolers were tested on batteries of theory of mind tasks, executive functioning tasks, and on their ability to adapt their lifting behavior to smoothly lift an unexpectedly heavy object. Results showed that children who adapted their lifting behavior more rapidly performed better on theory of mind tasks than those who adapted more slowly. These findings held up when age and performance on the executive functioning battery were statistically controlled. Although preliminary, we argue that this relation is attributable to individual differences in children's domain general abilities to efficiently change existing conceptual structures in response to experience.

1. Introduction

Between the ages of 3- and 5-years, children come to explicitly understand that the epistemic mental states that motivate observable behavior (e.g., belief, knowledge), are person-specific, idiosyncratic representations of reality (Perner, 1991). This representational theory of mind understanding (RTM) is often diagnosed with the "false belief" task (Wellman, Cross, & Watson, 2001; Wimmer & Perner, 1983) whereby children are asked to either predict or explain how a person will act when that person's beliefs do not match some true state of affairs. Scores of studies have shown that young 3-year-olds fail this task, and correct performance begins to develop around children's fourth birthday (see Wellman et al., 2001). However, much less work has been focused on the mechanisms underlying children's RTM development. In the present study, we used an individual differences approach to examine whether children's abilities to change their expectations about the weight of an object in a motor adaptation task is associated with preschoolers' theory of mind development.

Our focus on children's abilities to change their expectations stems from a consideration of one of the more prominent frameworks for investigating mechanisms that support preschoolers' development of an explicit RTM: the "theory theory" (Gopnik & Wellman, 1994; Wellman, 1990; Wellman & Gelman, 1998). The theory theory starts by emphasizing that mental states are theoretical constructs that can be used to generate expectations (i.e., hypotheses and predictions) about how people will act in particular situations. From a developmental standpoint, the question is not whether young children have some theory of mind; even young infants are presumed to have some understanding of the fact that behavior is motivated by internal mental states. These early understandings can be diagnosed in paradigms that investigate how infants react when expectations putatively generated by a consideration of internal mental states are violated (Onishi & Baillargeon, 2005; Phillips, Wellman, & Spelke, 2002; Woodward, 1998). Rather the developmental question is what kinds of expectations and explanations for behavior children's theories of mind tend to generate at different ages. Wellman and colleagues (Bartsch & Wellman, 1994; Wellman & Liu, 2004) have argued that children gradually change through a series of qualitatively different understandings of how mental states motivate behavior, finally arriving at a rudimentary but explicit adult-like theory of mind, sometime between children's third and fourth birthday.

For the present discussion, the most notable and perhaps also most controversial aspect of the theory theory is its proposal for how children's theories and expectations change. Change within theory theory is conceptualized as a process of adapting one's theories via a process akin to how formal scientists change the theories they use to explain empirical phenomena (Gopnik &

---

* Corresponding author.
E-mail address: sabbagh@queensu.ca (M.A. Sabbagh).

0893-6080/$ – see front matter © 2010 Elsevier Ltd. All rights reserved.
doi:10.1016/j.neunet.2010.08.007

Meltzoff, 1997; Gopnik & Wellman, 1992). As children navigate the world with one naïve theory, they may encounter instances in which their current theory leads to incorrect predictions or incoherent explanations for particular events. These experiences accumulate and spur children to adapt their theoretical constructs to achieve a better (i.e., more accurately predictive, more coherently and parsimoniously explanatory) understanding of how mental states relate to human behavior.

Evidence cited as generally consistent with theory theory comes from two literatures. First, 3-year-olds seem to have very different ways of explaining human actions than do 5-year-olds (Bartsch & Wellman, 1994). For instance, 3-year-olds tend to focus on the role that desire plays in explaining human action whereas 5-year-olds refer more to concepts of knowledge and belief (Bartsch, Campbell, & Troseth, 2007). This provides evidence suggesting that children do indeed go through qualitatively different phases of explaining how mental states relate to the world. Second, individual differences in experiential factors seem to predict the age at which preschoolers pass false belief tasks (Carpendale & Lewis, 2004). For instance, parent–child talk about mental states (Ruffman, Slade, & Crowe, 2002), number of siblings in the family (Ruffman, Perner, Naito, Parkin, & Clements, 1998), and socio-economic status (Pears & Moses, 2003) are all positively related to children's false belief development. These effects are expected if the processes that support changing their initial theories based upon their experiences are crucial for RTM development—those with more relevant experiences will change their theories more quickly than those with fewer relevant experiences (Bartsch, 2002).

Although these lines of evidence are largely consistent with theory theory, a number of authors have argued that both lines of evidence are also consistent with alternative theoretical perspectives (see e.g., Scholl & Leslie, 1999). Of course, no one line of evidence is likely to provide conclusive support for a broad theoretical framework such as the theory theory. One way of approaching this issue is to focus more squarely on the mechanisms that theory theorists propose to support change in theory of mind over the preschool years. At a glance, we might suggest that a suite of domain-general cognitive processes are involved in changing cognitive structures to better reflect experience. For one, to even see that change is necessary, one would need to notice prediction errors—instances in which a naïve expectation or hypothesis did not match what really happened. As an example in the theory of mind domain, if one thought that action was based primarily on desires (i.e., people do what they want to do), then viewing someone act on the basis of a false belief might result in a prediction error; there, someone wants to do something (i.e., find a stashed object) but acts in a way that does not straightforwardly comport with that desire (i.e., looks where the object is not). When prediction errors are made and identified as such, one might evaluate the quality and integrity of the new data, and then, based upon that analysis make a change to the existing conceptual structures responsible for generating the expectation. These changes would be made with the aim of making future expectations better match outcomes. It is important to note that this process of change is likely to be gradual rather than sudden. After all, a predictive system that changed too radically in response to any one piece of information would likely be too unstable for regular predictive power (e.g., Siegler & Chen, 1998). With respect to RTM development, it would seem that the transition from one theory of mind to another is gradual rather than abrupt, even when children are provided with a steady concentrated diet of rich information relevant to transitions in theory of mind (Amsterlaw & Wellman, 2006).

We know of no studies that have shown that these processes that enable cognitive change in response to experience are associated with RTM development, or any other transition in children's theory of mind reasoning. Part of this may be due to the somewhat protracted window within which changes in RTM occur (over roughly 18–24 months), and a general inability to parametrically assess how much impact a given experience has on the emergence of an RTM. Here, we propose that everyday phenomena in the domain of motor learning might provide a window on these processes. Several studies now have shown that people will lift a newly encountered object based upon prior expectations about the likely weight of the object. When lifting an object, people typically increase the vertical lift (or load) force to a target level that slightly exceeds the predicted weight of the object. When lifting a newly encountered object, weight predictions are based upon prior knowledge — or "internal models" — linking the size, material and identity of the object to weight (e.g., Flanagan & Beltzner, 2000; Flanagan, Bittner, & Johansson, 2008). The use of these predictions becomes manifest when people mis-lift (apply too much or too little force) objects because the object is unexpectedly light or heavy (Flanagan et al., 2008). Over repeated interactions with the object, people gradually (not suddenly) change their expectations about the force required to lift appropriately, and ultimately lift the object smoothly (Flanagan, Bowman, & Johansson, 2006).

What is perhaps most intriguing is that the processes that are thought to underpin load force adaptation (see e.g., Wolpert & Flanagan, 2009) broadly parallel those that theory theorists believe are important for spurring transitions in theory of mind. That is, both load force adaptation and theory change are thought to entail using a prior existing body of knowledge to generate a testable expectation about the world (either what an object will weigh or what a person will do). When prediction errors are made, the detection of the error promotes some adjustment to the system. The end result is that the systems that generate the expectations (either about object weight or about how mental states cause behavior) are revised to deliver more accurate expectations. Given these similarities of process, then, our main research question concerned whether children's abilities in a simple load–force adaptation paradigm might be associated with their RTM abilities. Finding such a relation could constitute evidence that domain general mechanisms that promote incremental change in conceptual structures are associated with RTM development.

In addition to this focal question, we included a battery of executive functioning (EF) tasks in our design as a potential control measure. Prior research has established that there is a connection between RTM and EF skills (see Benson & Sabbagh, 2009, for a recent review). In particular, RTM is associated with EF tasks that require children to inhibit a dominant or prepotent action or response in order to follow a rule that requires them to do something else. A number of researchers have noted that these "response-conflict" EF tasks require children to keep in mind two possible ways of acting on the world and select the one to engage based upon their awareness of the task context (Frye, Zelazo, & Burack, 1998). On the surface, it seems possible that similar processes may be at work when children encounter an unexpectedly heavy object; after becoming aware of the heavy object, children may need to recognize that the object can be lifted two ways and apply the force appropriate to the context. To determine whether this is the case, we included two well-established measures of children's response–conflict EF skills for inclusion in analyses.

2. Method

2.1. Participants

Thirty-three 42–54-month-old children (14 girls, $M_{\mathrm{age}} = 47.01$ months, $SD = 3.62$) were recruited to participate through a database of prospective participants recruited from a largely European–Canadian, middle-class professional and military community in Southeastern Ontario. Seven participants were excluded because of equipment malfunction during the load force adaptation task (see below). Thus, the final sample included 25 children (10 girls, $M_{\text{age}} = 47.2$ months, $SD = 3.60$). Participants' families were compensated for their participation.

2.2. Measures/materials

2.2.1. RTM battery

The RTM battery consisted of five tasks designed to assess children's understanding of epistemic mental representations (e.g., false belief). The RTM battery score was computed by summing the number tasks children successfully passed. The individual tasks, all of which have been used in previous research, are described briefly below along with scoring criteria.

Location false belief task (Wimmer & Perner, 1983). The experimenter first placed a doll-sized chest of drawers and a wood box in front on the child. Children were then introduced to two dolls, Andy and Heidi, playing with a plastic ball. The experimenter explained that after playing, Heidi had become tired and wanted to go take a nap. Before she left, Heidi put the ball away in the drawer. While Heidi was napping, Andy moved the ball from the drawer to the chest and then went out play. Heidi then returned, wanting to play with her ball. The experimenter then asked the child a test question (Where does Heidi think the ball is?) and two control questions to ensure that children remembered the events of the story (Where is Heidi put the ball? and Where is the ball now?). As is customary in individual differences designs, the questions were asked in the same order for each child for this and all of the tasks that we administered. Performance was coded as passing if children answered all three questions correctly.

Knowledge access false belief task (Wellman & Liu, 2004). In this task, children were shown a box and asked what they thought was inside. When children said that they did not know, the experimenter opened the box to reveal a plastic elephant. The box was then closed and the experimenter asked children to say again what is inside the box. After children answered, the experimenter introduced a tiger puppet and explained that tiger had never seen inside the box. The experimenter then asked children a test question (Does the tiger know what's in the box?) and a control question (Has the tiger ever seen inside the box?). Performance was coded as passing if children answered both questions correctly.

Contents false belief task (Gopnik & Astington, 1988). Children were shown a familiar candy box and was asked what was inside. After children answered with the name of the familiar candy, the experimenter showed that in fact there are crayons inside. The box was then closed and the experimenter introduced a monkey puppet who children were told had never seen inside the box. Next, the experimenter asked a test question ("What will the monkey think is in the box, candy or crayons?") and a control question ("Has the monkey ever seen inside the box"). Children who answered both questions correctly were coded as passing.

Appearance/reality task (Flavell, Green, & Flavell, 1986). There were two appearance/reality tasks. In the first, children were shown a sponge that was cut and painted to look like a rock. The experimenter first asked children what they thought the item looked like. After children responded "rock", they were shown that it was actually a sponge. Then children were asked a test question ("What does this look like, a rock or a sponge?") and a control question ("What is this really, a rock or a sponge?"). Children who answered both questions correctly were coded as passing.

In the second task, the experimenter showed children a picture of an orange castle that the experimenter then placed a blue plastic sheet over to make it look black. The experimenter covered and uncovered the castle several times and then finally asked a test question with the cover over ("When you look at this castle right now, does it look orange or black?") and a control question ("What color is the castle really and truly?"). Children who answered both questions correctly were coded as passing.

2.2.2. Executive functioning battery

The executive functioning battery included two tasks, both of which have been used in prior research, and are described in brief below. The scores of these two tasks were standardized and averaged to create an executive functioning composite score.

Grass–snow stroop (Carlson & Moses, 2001). Children were presented with a black board with a white card fixed to the top right corner and a green card fixed to the top left corner. On the bottom of the board, centered, were two felt hand shapes for children's hands. After ensuring that children knew the colors associated with grass (green) and snow (white), the experimenter explained that they were going to play a 'silly game' in which children were to "point to white when I say grass, and green when I say snow". Practice trials with feedback were given until children responded correctly to each of the 'grass' and the 'snow' prompts. For the test trials, the experimenter read out 16 prompts in the following order: Grass, Snow, Snow, Grass, Snow, Grass, Grass, Snow, Snow, Grass, Snow, Grass, Grass, Snow, Grass, and Snow. No feedback was given on test trials. The score on the task was simply the number of correct responses over the 16 trials. If multiple responses were given, only the first response was coded.

Dimensional-change card sort (Zelazo, 2006). This task required children to sort a deck of cards first according to their shape, and then according to their color. Cards depicted red rabbits, blue rabbits, red boats and blue boats. Children were presented with two baskets, one labeled with a red rabbit (the left basket) and the other labeled with a blue boat (the right basket). The experimenter first explained the task as the 'shape game' in rabbits were sorted into the rabbit basket and boats into the boat basket. After providing an example of each shape card being sorted, children were presented with five cards, one at a time, and asked into which basket the card should be sorted. Then, the experimenter switched the game to the 'colour game' in which cards all blue cards were to be sorted into the blue boat basket, and red cards into the red rabbit basket. There are five cards to be sorted; two could be sorted correctly in a way that was compatible with the old rule (i.e., red rabbits and blue boats) whereas the other three were incompatible and thus more diagnostic of following the new rule. Children were reminded of the rule before each card was presented but not given any corrective feedback about their performance. The dependent measure was the number of incompatible cards that children sorted correctly according to the new rule (0–3).

2.2.3. Load force adaptation task

Previous work on object lifting in children has shown that the ability to scale load force for object size emerges at 3.5 years of age (Gordon, Forssberg, Johansson, Eliasson, & Westling, 1992) and the ability to adapt load forces based on previous lifts of the object emerges a year earlier (Forssberg, Eliasson, Kinoshita, Johansson, & Westling, 1991). In the latter study, the weight of a single test object was expectedly switched between trials and the results showed that 2.5 year olds' load force was influenced by the weight in the previous lift. However, changes in load force over repeated trials were not examined nor were children's responses to unexpected changes in weight. Thus, we created a novel task to study load–force adaptation in preschoolers.

Children were first shown a green copper cube (4.3 × 4.3 × 4.3 cm) that was hollow and weighed 3 N. Attached to the top of the cube was a small clear 2.5 cm plastic tab to facilitate lifting.

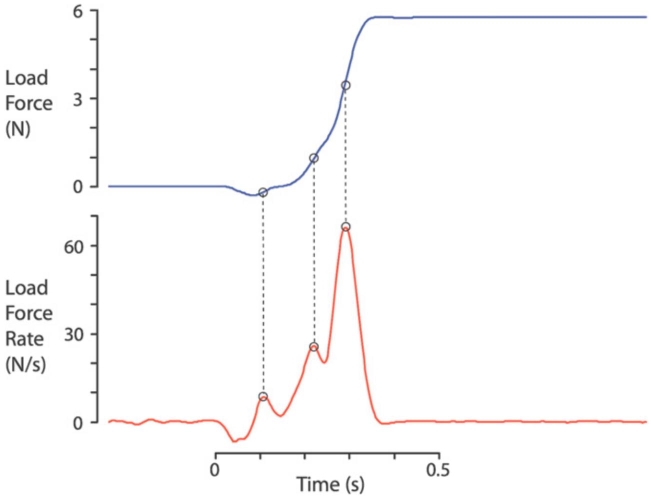

Fig. 1. Displays of (a) load–force and (b) rate of change of load force as a function of time for a single lift of a 6 N cube. Dots on the graphs represent the time points for the algorithmically defined rate-of-change peaks occurring throughout the lift.

The cube was placed on a force sensor (Nano 17 F/T sensor, ATI Industrial Automation, Inc., Garner, NC, USA) that stood 3.2 cm high. Children were simply asked to lift the cube off the sensor by the plastic tab and onto a patch of blue construction paper (2.5 × 2.5 cm) affixed to a small shelf that stood 10.2 cm high off the table just to the right or left of the sensor (depending on the child's handedness). The height of the lift from the sensor to the shelf was 6.4 cm. Children were asked to repeat this eight times, and each lift was counted aloud by the experimenter as it was completed. After the 8 lifts the 3 N cube was taken away and, in view of the children, replaced with a copper cube that looked the same (same color, material, and dimensions), but was filled and thus weighed 6 N. We reasoned that children, having just become accustomed to the 3 N cube, would find the new cube unexpectedly heavy and initially not apply sufficient force to smoothly complete the lift. Thus, we expected children to use their experience to overturn their initial hypotheses about the weight of the cube and adapt their load force over successive lifts. Children were asked to lift the new 6 N cube eight times ("Now let's try this one!"), again with each lift counted aloud as it was completed. Children were given encouraging feedback after each lift (e.g., "okay", "thanks", etc.,). Any comments children made about the differing weights of the object were ignored by the experimenter who simply kept providing encouraging feedback.

As participants lifted the objects from the sensor to the shelf, the sensor recorded the vertical load at a rate of 295 Hz. The raw data for each lift were imported into MATLAB (MathWorks, Natick, MA), smoothed using a low-pass butterworth filter with a 20 Hz cut-off, and then differentiated to obtain the rate of change of load force with respect to time (see Fig. 1). Previous research has shown that people lift familiar objects by smoothly increasing the load force to a target level that slightly exceeds the expected weight of the object. This smooth increase in load force is characterized by an approximately bell-shaped rate of change of load force such that the peak rate of change of load force occurs when the load force is about half the expected weight of the object. Thus, how much a person *expects* an object to weigh can be estimated by how much load force is being exerted on the object at that initial peak in the rate of change. For instance, when people expect an object to weigh 6 N, the rate-of-change peak occurs when they are exerting approximately a 3 N force on the object.

A purpose-designed script was written to identify the load force present at each clear rate-of-change peak (algorithmically identified) that occurred over the course of all lifts for all participants. These data were then visually inspected by a coder who identified the initial rate-of-change peak for further statistical analysis. The initial rate-of-change peak was defined as the first algorithmically identified rate-of-change peak that was associated with a force of greater than 0.5 N. Peaks early in the lift that were associated with forces of less than 0.5 N were not attributable to obvious attempts to lift the object. Coders could not be blind to the object that was being lifted because that was apparent from the total load force applied to execute the lift. However, coders were blind to children's age and their performance on the other tasks. Two coders coded all of the trials and interrater agreement was 89%. All discrepancies were resolved through discussion.

2.3. Procedure and design

All children were tested in an on-campus playroom at a child-sized table with a single experimenter. After a brief warm-up phase in which children played with assorted toys unrelated to items used in the present study, children gave their assent to participate in the study. Participants' guardians watched the entire procedure on a closed-circuit television from another room.

Because this is an individual differences study, all tasks were presented the same way to each child, in the same fixed order (see e.g., Carlson & Moses, 2001): (1) Load force adaptation, (2) Location Change False Belief Task, (3) Knowledge Access False Belief Task, (4) Grass–Snow Stroop, (5) Dimensional Change Card Sort, (6) Contents False Belief Task and (7) Appearance/Reality Task. The entire session took about 25 min.

3. Results

The number of participants passing each task and the descriptive statistics for the RTM and executive functioning batteries as a whole are presented in Table 1. The proportion of children passing each RTM task in the present study is consistent with findings of other studies that have used similar aged children (see e.g., Sabbagh, Moses, & Shiverick, 2006), as was the executive functioning data (see Carlson, 2005). Close inspection of the distribution for the RTM battery totals revealed a somewhat positively skewed distribution, with many children passing either none or just one of the tasks. This was also expected given the age of our participants. Thus, for our main analyses we dichotomized children's RTM battery performance based upon a median split: low (children who passed 0–1 tasks, $n = 11$) and high (children who passed 2 or more tasks, $n = 14$). Children in the high RTM group were older ($M = 48.60$ months, $SD = 3.52$) than children in the low RTM group ($M = 45.45$ months, $SD = 3.05$), $t(23) = 2.35$, $p = 0.028$. Children's RTM performance was not significantly correlated with performance on the EF battery, $r(23) = 0.324$, $p = 0.114$, though, the magnitude of the relation we observed is similar to that seen in previous studies that have investigated the relation (e.g., Sabbagh et al., 2006).

With respect to the load–force adaptation task, the average force applied at the initial rate-of-change peak for all subjects and all lift trials are presented in Fig. 2. There are two notable aspects of this figure. For one, with respect to the lighter 3 N cube, it appeared that children had to learn the weight of this novel object. At first, children were only applying about 1 N of force at the initial rate-of-change peak, which is slightly under the 1.5 N of force that would be expected for its weight. Over the course of the first 4 trials, however, children on average adapted their force and eventually applied 1.5 N of force at the initial rate-of-change peak, and continued to do so for the remainder of the lifts with the 3 N cube. Second, with respect to the heavier 6 N cube, it was clear that children initially expected this to be the same weight as the 3 N cube—children only applied 1.5 N at the initial rate-of-change peak

Table 1
Summary of performance on RTM and executive functioning task batteries.

Task	Correct
RTM battery[a]	
Location false belief	28% (7)
Contents false belief	24% (6)
Knowledge access	40% (10)
Appearance/reality: rock/sponge	40% (10)
Appearance/reality: castle	60% (15)
Battery total (out of 5)	$M = 1.88$ ($SD = 1.394$)
Executive functioning battery	
Dimensional change card sort (out of 3)	$M = 1.791$ ($SD = 1.353$)
Grass/snow stroop (out of 16)	$M = 7.6$ ($SD = 5.377$)

Notes: RTM = Representational theory of mind.
[a] For individual RTM tasks, figures represent percent (and number) of children who passed the task.

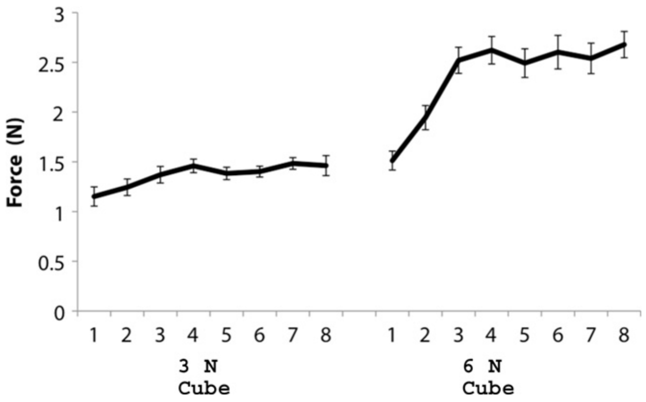

Fig. 2. Mean load force (N) (+/− 1 SE) at initial rate-of-change peaks for all 8 trials for the 3 and 6 N cubes, collapsed across RTM groups.

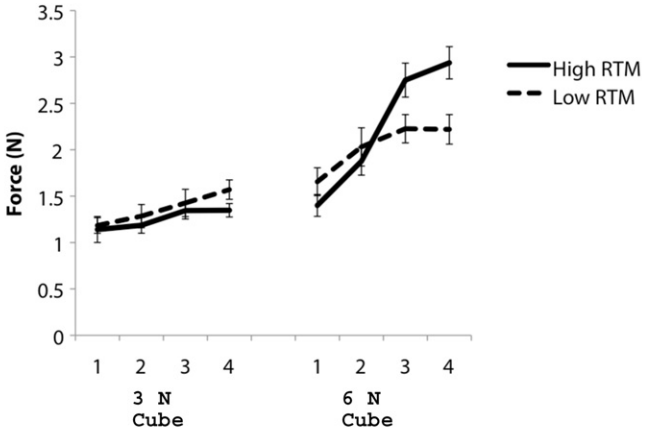

Fig. 3. Mean load force (N) (+/− 1 SE) at initial rate-of-change peak for each of the first four lifting trials for high and low RTM groups.

for the first lift of the object. Again, children gradually adapted their load force over the first four trials with the 6 N object, and load force at the initial rate-of-change peak approached the expected level of 3 N. These findings show that as a group, children adapted load force to object weight based on experience from previous lifts, and that the end product of that adaptation was to apply load force at initial rate-of-change peaks in roughly the same way an adult would.

The focal question was whether children in the two RTM groups differed in their adaptation patterns. Because the trajectories of change in force applied at the initial rate-of-change peak were clear and appeared to level off after the first four lifts with each object, our main analyses focused on these trials. Collapsed across groups, the overall mean trajectory across the first four trials for each object was described very well by a linear function (light object: linear $r^2 = 0.973$, heavy object: linear $r^2 = 0.940$). Thus, to derive an adaptation measure for each participant and each object, we fit a linear function to the force at initial rate-of-change peak for the first four lifts of each cube and calculated the slope coefficient. We reasoned that the slope coefficient would be an appropriate adaptation measure as this captures the magnitude and direction of change in each participant's force at initial rate-of-change peak across the four lifts (i.e., larger slope coefficients indicate more rapid increases in force).

The average force at initial rate-of-change peak for the first four trials of each cube for low and high RTM groups are presented in Fig. 3. Inspection of this figure shows that although the two groups appeared to adapt at similar rates to the initial 3 N cube, the high RTM group appeared to adapt more rapidly than the low RTM group to the 6 N cube. To characterize this effect statistically, we conducted a 2 (Cube: 3 N vs. 6 N cube) × 2 (RTM group) mixed model ANOVA with Cube entered as a within-subjects factor, RTM group as a between-subjects factor, and adaptation slope as the dependent measure. This analysis revealed a significant main effect of cube, $F(1, 23) = 18.393, p < 0.001$, partial $\eta^2 = 0.444$, whereby collapsed across RTM group the average adaptation slope of the 6 N cube ($M = 0.369$, $SD = 0.053$) was steeper than the average adaptation slope of the 3 N cube ($M = 0.108$, $SD = 0.040$). However, this main effect was qualified by a significant cube by RTM group interaction, $F(1, 23) = 14.741, p = 0.001$, partial $\eta^2 = 0.391$.

Follow-up t-tests were calculated to investigate the source of the significant interaction between adaptation slope and RTM group. For the slopes of the 3 N cube, the difference between the low RTM group ($M = 0.162$, $SD = 0.171$) and the high RTM group ($M = 0.053$, $SD = 0.215$) was not significant, $t(23) = 1.36, p = 0.187$. In contrast, for the slopes of the 6 N cube, the difference between the RTM groups was significant, $t(23) = 3.363, p = 0.003$, partial $\eta^2 = 0.330$. The high RTM group ($M = 0.549$, $SD = 0.246$) had a significantly steeper adaptation slope than the low RTM group ($M = 0.189$, $SD = 0.289$).

The fact that age and EF were either significant or near-significant predictors of RTM group, a crucial question is whether these factors might account for the relation between RTM and adaptation to the 6 N object. Preliminary analyses showed that neither EF nor age were significant predictors of the 6 N adaptation slope, $rs(23) = -0.058$ and 0.279, $ps = 0.782$ and 0.176 respectively. The fact that neither variable is associated with slopes makes it unlikely that the relation between RTM and 6 N adaptation slopes is mediated by EF or age. Nonetheless, to better understand the nature of these relations, we conducted a binary logistic regression in which RTM group membership was predicted from the adaptation slopes for the 6 N cube, while statistically controlling for children's age and executive functioning performance. All predictor variables were standardized to render comparable odds ratios from the regression. The test showed that when all three predictors were entered into the regression, the 6 N adaptation slope was the strongest and only significant predictor of RTM group membership, Wald = 4.47, $p = 0.034$. Odds ratio analysis showed that children with adaptations slopes greater than 1 standard deviation above the mean were 6.9 times more likely to also be in the high RTM group. Age and executive functioning did not meet standard significance levels (Wald = 2.10, $p = 0.147$ and Wald = 2.343, $p = 0.126$, respectively), though, as in the uncontrolled analyses, the trends were in the expected direction.

To finish, we conducted two analyses to assess possible artifactual reasons for our findings. For one, we noted that one reason the slopes may have been weaker in the low versus high RTM group was that they seemed to exert slightly more force at initial rate-of-change peak on the first trial of the 6 N cube. Although this difference was not statistically significant, $t(23) = 1.35, p = 0.187$, the slightly raised intercept for the RTM group may have weakened their slope because they started closer to the optimal level of force. Indeed, for the entire sample, initial force at first lift was negatively correlated with slope, $r(23) = -0.474, p = 0.017$. To ensure that this did not account for our findings, we residualized the adaptation slope measure controlling for force at initial peak rate-of-change in the 6 N cube, and substituted this as the dependent measure in the analyses described above. All patterns of significance remained the same, and were of roughly equal strength, thereby giving us confidence that it was adaptation *per se* (and not an artifact of force at first lift) that accounts for the effects.

Second, we noted that the load force adaptation task bears some surface similarities to the appearance–reality tasks that were included in our RTM battery. In the appearance/reality paradigm, an object is shown to the child which appears to be one thing though it is really something different. The same is true for our load force adaptation task; the second cube appears that it would weigh 3 N (based upon prior experience) but it truly weighs 6 N. Perhaps the association between load force adaptation and RTM performance is due in part to this similarity. Although there are important differences in the patterns of findings between the two tasks (which we will discuss below), we wanted to ensure that the relation was not accounted for by the similarities of the tasks. Thus, we reran the focal analyses creating the high and low RTM groups without the appearance–reality tasks. The pattern of findings was unchanged from the full scale analyses, thereby providing confidence that the relation between RTM performance and load force adaptation is not attributable to similarities between the load force adaptation task and the appearance–reality tasks.

4. Discussion

The goal of the present study was to explore whether domain general processes underlying adaptation are associated with preschool children's developing representational theory of mind. To this end, we suggested that characterizing the manner in which children adapt their lifting behavior when presented with an unexpectedly heavy object might provide a measure of children's cognitive adaptation skills. This measure was used to determine whether individual differences in adaptation were associated with preschoolers' RTM skills. We found that indeed, individual differences in load–force adaptation predicted whether children showed strong or poor performance on a battery of RTM tasks. The relation between RTM and load–force adaptation was statistically significant and strong even when executive functioning and age were statistically controlled. These findings provide some confidence that the relation between RTM and load–force adaptation cannot be attributable to broad maturational factors, or some specific cognitive developmental factors (e.g., conflict monitoring, response selection) that share surface similarities with RTM and load–force adaptation.

Our initial motivation for conducting this preliminary investigation was the prominent "theory theory" framework that has been proposed as an account of the mechanisms underlying transitions in children's theory of mind reasoning, including preschoolers' transition to a sophisticated RTM (Gopnik & Wellman, 1994). Briefly, theory theorists have suggested that children change their theories of mind as they come to recognize that their naïve theories lead to incorrect predictions of human behavior. Detecting these errors is thought to spur processes of adaptation whereby children make gradual changes to a current naïve theory that ultimately coalesce in a new, more accurate theory. Similar processes, albeit on a much shorter time scale, have been proposed to account for the processes by which people adapt their lifting behavior when encountering an unexpectedly heavy or light object (e.g., Flanagan et al., 2006). By showing that children's RTM performance is associated with the speed with which they adapted to an unexpectedly heavy load, we provide support for the notion that individual differences in children's ability to efficiently change existing conceptual structures in response to experience is an important factor in RTM development. In turn, this pattern of results provides some support for the theory theory approach to RTM development more generally.

Yet, an important question concerns whether the load force task indeed indexes individual differences in preschoolers' abilities to change existing ideas about the world. An alternative possibility is that the task measures incremental learning more broadly. We think, however, that the fact that RTM reasoning was associated with adaptation slopes for the 6 N cube, but not the 3 N cube, provides some evidence against this alternative possibility. Smoothly lifting the initial 3 N cube required learning. As is apparent from the mean trajectories, children on average applied approximately 1 N of force at the initial rate-of-change peak thereby suggesting that children initially thought the cube would weigh 2 N. Children then had to learn over the first few trials that the cube in fact weighed 3 N. The 3 N cube, however, was novel to children. Thus children's prior hypotheses about the weight of the cube were likely untested and weak. In the case of the 6 N cube, children likely had a stronger hypothesis as to how much the cube should weigh because of their immediately preceding experience with an identical looking cube. Thus, changing force in the 6 N cube likely relies to a greater extent on the processes that are associated with changing an established, previously well-supported hypothesis about the world. This pattern of findings provides a basis for suggesting that preschool children's RTM skills are not associated with learning in general, but rather with the processes involved in changing an established idea in the face of clear evidence to the contrary.

A second challenge to our preferred interpretation is a more theoretical one. Although on their surfaces, RTM and load–force adaptation would appear to have very little in common, they may have some similar representational requirements (see e.g., Perner, 1991). In order to accurately pass an RTM task, one must keep in mind two competing representations (i.e., beliefs) of the same objective situation. In some sense, the load force adaptation task shares this characteristic; that is, adaptation could be facilitated if children can entertain the idea that the identical green cubes can either weigh 3 or 6 N. It could be that children who have the representational capacity to explicitly reflect on the possibility that two identical cubes can differ in weight are better able to rapidly adapt their lifting behavior. This ability to make a "many-to-one" mapping has been proposed to account for associations between RTM and children's understanding of certain linguistic phenomena including homonyms (i.e., cases in which the same word has two different referents) (Doherty & Perner, 1998) and non-literal language understanding such as metaphor (i.e., cases in which the same utterance can have alternate meanings) (Happe, 1993).

It is not possible for us to definitively rule out this alternative interpretation with the present data. However, it is worth noting that there are important differences between the pattern of findings in the load–force adaptation paradigm and the typical errors that children made in representational tasks such as the false belief and appearance–reality tasks. For example, the standard error in the appearance–reality task such as the sponge–rock task is that after learning the object is a sponge, children then say that it also *looks* like a sponge, though a moment ago they said it looked like a rock (Flavell et al., 1986). Thus, the new information about reality supplants the representation that was based on prior learning (i.e., all of their answers are colored by a "reality error").

The error children appear to make in the load force adaptation task is quite the opposite, though. Although it is clear that the new cube in reality weighs 6 and not 3 Ns, children in the low RTM group appear to be slower to take what they have learned about reality to overcome their initial guesses. Given that the pattern of performance in the low RTM group seems to reflect slower adaptation rather than a common "reality error" we believe that the association between the load–force adaptation task and the RTM battery is currently best explained by the fact that both appear to rely on processes of conceptual change in response to experience.

Additional lines of inquiry provide some basis for making increasingly specific speculations about the processes of conceptual change that are shared between RTM development and load–force adaptation. For instance, recent theoretical work has been done connecting the processes of conceptual change within the theory theory framework to computational models that rely on Bayesian learning models (Gopnik & Tenenbaum, 2007; Tenenbaum, Griffiths, & Kemp, 2006). In brief, Bayesian learning models take as their starting place that a given predictive system generates expectations about what the world will be like based upon prior hypotheses of variable strength regarding some relevant causal model. When experience provides relevant data, the strength of this prior hypothesis is adapted based upon using what is called the "posterior probability" computation, which is based upon combining some prior probability of the hypothesis being true with an assessment of the likelihood of observing the data independent of the hypothesis being true, and the probability that the data would also be observed under some alternative hypotheses (Gopnik et al., 2004). Using these computations, instances of data that are uniquely consistent with the prior hypothesis gradually strengthen the hypothesis (and weaken alternatives). Conversely, data that are inconsistent weaken the prior more or less depending on how consistent the data are with the alternatives. A growing literature has demonstrated that young children's causal learning of new information relies on Bayesian processes (Kushnir & Gopnik, 2005; Sobel, Tenenbaum, & Gopnik, 2004; Xu & Tenenbaum, 2007). Most interesting with respect to the present context is that Bayesian processes have been postulated to account for people's performance in load–force adaptation tasks (see Flanagan et al., 2006). Together, these two lines of research promote a specific speculation that load–force adaptation and RTM development may be linked because both rely on the kinds of Bayesian computations that have been proposed to promote conceptual change in response to experience.

A second specific mechanism that may connect the processes of conceptual change in load–force adaptation and RTM is the neurotransmitter dopamine (DA). Basic research in behavioral neuroscience has shown that dopaminergic activity is elicited when animals encounter situations in which their expectations about an event (such as a reward) do not match with what ultimately occurs (Schultz, 2000). It is generally thought that in these cases, dopaminergic activity promotes plasticity necessary for adjusting expectations, and coming to increasingly refined understandings of the causal structure of a given event (Schultz, 2007). Individuals with Parkinson's disease, which is associated with deficits in DA function, have difficulty with tasks that require sensori–motor adaptation (e.g., Krebs, Hogan, Hening, Adamovich, & Poizner, 2001). More recent work has shown that the difficulties experienced by individuals with Parkinson's Disease are directly attributable to disturbances in DA functioning (Paquet et al., 2008). Although we know of no work that has directly related DA functioning to load–force adaptation, it seems likely that such a relation exists given that load–force adaptation is a kind of sensori–motor adaptation. Recent research has, however, shown that individual differences in DA functioning are associated with preschoolers' RTM development (Lackner, Bowman, & Sabbagh, 2010). Together, the animal, clinical, and developmental evidence provide a basis for the speculation that the conceptual updating processes that are proposed to be critical for both load–force adaptation and RTM development may have a neurobiological basis in DA functioning.

Finally, there is limited evidence that RTM reasoning and load–force adaptation may rely on partially shared neural substrates. The neural bases of RTM reasoning in adults has been well-researched over the past 10 years, with consensus building around the conclusion that RTM relies on a relatively circumscribed network of brain areas including the medial prefrontal cortex and the right temporal parietal juncture (e.g., Amodio & Frith, 2006; Saxe, 2006). Recent research has shown that the same is true for preschool-aged children (Sabbagh, Bowman, Evraire, & Ito, 2009). Although, research on the neural systems that are engaged during load–force adaptation is sparse, one recent study suggests that the right temporal–parietal juncture plays a critical role (Jenmalm, Schmitz, Forssberg, & Ehrsson, 2006). Of particular interest in this study was the finding that the right temporal parietal juncture was active anytime there was a prediction–outcome mismatch regarding an object's weight. The researchers suggested that this finding is evidence that the right temporal parietal juncture is involved in the process of updating prior hypotheses about weight towards more accurate predictions. Given the putative role of this same process in RTM development, we might speculate that the right temporal parietal juncture provides a neuroanatomical substrate for conceptual updating processes that are common to both load–force adaptation and RTM development.

There are important limitations to the present study. First, our sample size is relatively small. Small sample sizes are typically a problem for statistical power, and the fact that our findings were strong in the face of a small sample size likely speaks to the fact that we observed a robust relation between load–force adaptation and RTM development. Small sample sizes do, however, raise questions about whether these findings would generalize. Thus, replication of the present study with a larger sample size would help to further support the claims we have made here.

A second limitation is potentially more serious. As outlined above, our main dependent measure of load–force adaptation was the slope of the change in force applied at the initial rate-of-change peak for the first four trials of the 6 N object. Stronger support for the claim that children's performance with the 6 N object was attributable to adaptation and conceptual updating per se would come from a comparison with participants who lifted the 6 N cube *prior* to the 3 N cube. Without this comparison, we cannot tell conclusively whether children in the high RTM group are better at changing strong prior hypotheses about weight (attributable to their prior lifting of the 3 N cube) or simply better at learning how to lift a 6 N cube. This comparison was not possible in the present individual differences design in which it is customary to provide all children with the same experience so that artifact (fatigue, item effects, etc.) is the same across children. Of course, we know of no theoretical reason that simply learning how to lift a 6 N cube, but not a 3 N cube, would show an empirical association with RTM performance. Nonetheless, running such a condition in a more experimental paradigm would provide important additional support for our preferred interpretations.

To summarize, we found that children's performance in a load–force adaptation task was associated with their performance on an RTM battery. Our preferred hypothesis is that this relation is attributable to the fact that both load–force adaptation and RTM development rely on a common process of revising prior existing conceptual structures in response to experience of prediction–outcome mismatch. There are computational, neurochemical and neuroanatomical mechanisms that have been proposed to support such a revision process, and prior literature has linked each mechanism to RTM reasoning and to load–force or sensori–motor adaptation. More broadly, these findings provide support for the view that children's RTM development is best characterized as a conceptual development that emerges as domain-general cognitive processes interact with domain-specific experiences and conceptual structures.

Acknowledgements

This research was supported by Discovery Grants from the Natural Sciences and Engineering Research Council of Canada (NSERC) to Sabbagh and to Flanagan. We thank Taryn Ridsdale, Victoria Nanowski, Deborah Grossman and Tasha Hanuschak for their help with data collection and data coding. We thank Valerie Kuhlmeier for helpful comments and discussion of the manuscript.

References

Amodio, D. M., & Frith, C. D. (2006). Meeting of minds: the medial frontal cortex and social cognition. *Nature Reviews Neuroscience*, *7*, 268–277.

Amsterlaw, J., & Wellman, H. M. (2006). Theories of mind in transition: a microgenetic study of the development of false belief understanding. *Journal of Cognition and Development*, *7*, 139–172.

Bartsch, K. (2002). The role of experience in children's developing folk epistemology: review and analysis from the theory–theory perspective. *New Ideas in Psychology*, *20*, 145–161.

Bartsch, K., Campbell, M. D., & Troseth, G. L. (2007). Why else does Jenny run? Young children's extended psychological explanations. *Journal of Cognition and Development*, *8*, 33–61.

Bartsch, K., & Wellman, H. M. (1994). *Children talk about the mind*. Cambridge: Cambridge University Press.

Benson, J. E., & Sabbagh, M. A. (2009). Theory of mind and executive functioning: a developmental neuropsychological approach. In P. D. Zelazo, M. Chandler, & E. Crone (Eds.), *Developmental social cognitive neuroscience* (pp. 63–80). New York: Psychology Press.

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. (2005). Developmentally sensitive measures of executive function in preschool children. *Developmental Neuropsychology*, *28*, 595–616.

Carpendale, J. I. M., & Lewis, C. (2004). Constructing an understanding of mind: the development of children's social understanding within social interaction. *Behavioral and Brain Sciences*, *27*, 79–150.

Doherty, M., & Perner, J. (1998). Metalinguistic awareness and theory of mind: just two words for the same thing? *Cognitive Development*, *13*, 279–305.

Flanagan, J. R., & Beltzner, M. A. (2000). Independence of perceptual and sensorimotor predictions in the size–weight illusion. *Nature Neuroscience*, *3*, 737–741.

Flanagan, J. R., Bittner, J. P., & Johansson, R. S. (2008). Experience can change distinct size–weight priors engaged when lifting objects and judging their weights. *Current Biology*, *18*, 1742–1747.

Flanagan, J. R., Bowman, M. C., & Johansson, R. S. (2006). Control strategies in object manipulation tasks. *Current Opinion in Neurobiology*, *16*, 1–10.

Flavell, J. H., Green, F. L., & Flavell, E. R. (1986). *Monographs of the society for research in child development*: *Vol. 51*. *Development of knowledge about the appearance–reality distinction* (Serial No. 212).

Forssberg, H., Eliasson, A. C., Kinoshita, H., Johansson, R. S., & Westling, G. (1991). Development of human precision grip 1: basic coordination of force. *Experimental Brain Research*, *85*, 451–457.

Frye, D., Zelazo, P. D., & Burack, J. A. (1998). Cognitive complexity and control: I. Theory of mind in typical and atypical development. *Current Directions in Psychological Science*, *7*, 116–121.

Gopnik, A., & Astington, J. W. (1988). Children's understanding of representational change and its relation to the understanding of false belief and the appearance–reality distinction. *Child Development*, *59*, 26–37.

Gopnik, A., Glymour, C., Sobel, D. M., Schulz, L. E., Kushnir, T., & Danks, D. (2004). A theory of causal learning in children: causal maps and Bayes nets. *Psychological Review*, *111*, 3–32.

Gopnik, A., & Meltzoff, A. (1997). *Words, thoughts, and theories*. Cambridge, MA: MIT Press.

Gopnik, A., & Tenenbaum, J. B. (2007). Bayesian networks, Bayesian learning and cognitive development. *Developmental Science*, *10*, 281–287.

Gopnik, A., & Wellman, H. M. (1992). Why the child's theory of mind really is a theory. *Mind & Language*, *7*, 145–169.

Gopnik, A., & Wellman, H. M. (1994). The theory theory. In L. A. Hirschfeld, & S. A. Gelman (Eds.), *Mapping the mind: domain specificity in cognition and culture*. New York: Cambridge University Press.

Gordon, A. M., Forssberg, H., Johansson, R. S., Eliasson, A. C., & Westling, G. (1992). Development of human precision grip. III. Integration of visual size cues during the programming of isometric forces. *Experimental Brain Research*, *90*, 399–403.

Happe, F. (1993). Communicative competence and theory of mind in autism: a test of relevance theory. *Cognition*, *48*, 101–119.

Jenmalm, P., Schmitz, C., Forssberg, H., & Ehrsson, H. H. (2006). Lighter or heavier than predicted: neural correlates of corrective mechanisms during erroneously programmed lifts. *Journal of Neuroscience*, *26*, 9015–9021.

Krebs, H. I., Hogan, N., Hening, W., Adamovich, S. V., & Poizner, H. (2001). Procedural motor learning in Parkinson's disease. *Brain Research*, *141*, 425–437.

Kushnir, T., & Gopnik, A. (2005). Children infer causal strength from probabilities and interventions. *Psychological Science*, *16*, 678–683.

Lackner, C. L., Bowman, L. C., & Sabbagh, M. A. (2010). Dopaminergic functioning and preschoolers' theory of mind. *Neuropsychologia*, *48*, 1767–1774.

Onishi, K., & Baillargeon, R. (2005). Do 15-month-old infants understand false beliefs? *Science*, *308*, 255–258.

Paquet, F., Bedard, M. A., Levesque, M., Tremblay, P. L., Lemay, M., Blanchet, P. J., et al. (2008). Sensorimotor adaptation in Parkinson's disease: evidence for a dopamine dependent remapping disturbance. *Experimental Brain Research*, *185*, 227–236.

Pears, K. C., & Moses, L. J. (2003). Demographics, parenting and theory of mind in preschool children. *Social Development*, *12*, 1–20.

Perner, J. (1991). *Understanding the representational mind*. Cambridge, MA: MIT Press.

Phillips, A. T., Wellman, H. M., & Spelke, E. S. (2002). Infants' ability to connect gaze and emotional expression to intentional action. *Cognition*, *85*, 53–78.

Ruffman, T., Perner, J., Naito, M., Parkin, L., & Clements, W. (1998). Older (but not younger) siblings facilitate false belief understanding. *Developmental Psychology*, *34*, 161–174.

Ruffman, T., Slade, L., & Crowe, E. (2002). The relation between children's and mothers' mental state language and theory-of-mind understanding. *Child Development*, *73*, 734–751.

Sabbagh, M. A., Bowman, L. C., Evraire, L. E., & Ito, J. M. B. (2009). Neurodevelopmental correlates of theory of mind in preschool children. *Child Development*, *80*, 1147–1162.

Sabbagh, M. A., Moses, L. J., & Shiverick, S. M. (2006). Executive functioning and preschoolers' understanding of false beliefs, false photographs and false signs. *Child Development*, *77*, 1034–1049.

Saxe, R. (2006). Uniquely human social cognition. *Current Opinion in Neurobiology*, *16*, 235–239.

Scholl, B. J., & Leslie, A. M. (1999). Modularity, development, and 'theory of mind'. *Mind & Language*, *14*, 131–153.

Schultz, W. (2000). Multiple reward signals in the brain. *Nature Reviews Neuroscience*, *1*, 199–207.

Schultz, W. (2007). Behavioral dopamine signals. *Trends in Neurosciences*, *30*, 203–210.

Siegler, R. S., & Chen, Z. (1998). Developmental differences in rule learning: a microgenetic analysis. *Cognitive Psychology*, *36*, 273–310.

Sobel, D. M., Tenenbaum, J. B., & Gopnik, A. (2004). Children's causal inferences from indirect evidence: backwards blocking and Bayesian reasoning in preschoolers. *Cognitive Science*, *28*, 303–333.

Tenenbaum, J. B., Griffiths, T. L., & Kemp, C. (2006). Theory-based Bayesian models of inductive learning and reasoning. *Trends in Cognitive Sciences*, *10*, 309–318.

Wellman, H. M. (1990). *The child's theory of mind*. Cambridge, MA: MIT Press.

Wellman, H. M., Cross, D., & Watson, J. (2001). Meta-analysis of theory of mind development: the truth about false-belief. *Child Development*, *72*, 655–684.

Wellman, H. M., & Gelman, S. A. (1998). Knowledge acquisition in foundational domains. In P. Mussen, W. Damon, & R. S. Siegler (Eds.), *Cognition, perception, and language*: *Vol. 2*. *Handbook of child psychology* (5th ed.) (pp. 523–572). New York: Wiley.

Wellman, H. M., & Liu, D. (2004). Scaling of theory of mind tasks. *Child Development*, *75*, 523–541.

Wimmer, H., & Perner, J. (1983). Beliefs about beliefs: representation and constraining function of wrong beliefs in young children's understanding of deception. *Cognition*, *13*, 103–128.

Wolpert, D. M., & Flanagan, J. R. (2009). Forward models. In T. Baynes, A. Cleeremans, & P. Wilken (Eds.), *Oxford companion to consciousness*. Oxford: Oxford University Press.

Woodward, A. L. (1998). Infants selectively encode the goal object of an actor's reach. *Cognition*, *69*, 1–34.

Xu, F., & Tenenbaum, J. B. (2007). Sensitivity to sampling in Bayesian word learning. *Developmental Science*, *10*, 288–297.

Zelazo, P. D. (2006). The dimensional change card sort (DCCS): a method of assessing executive function in children. *Nature Protocols*, *1*, 297–301.