Research Article

Psychological Science
2014, Vol. 25(9) 1712–1721

Reprints and permissions:
sagepub.com/journalsPermissions.nav
DOI: 10.1177/0956797614538458
pss.sagepub.com

Young Children Bet on Their Numerical Skills: Metacognition in the Numerical Domain

Vy A. Vo[1], Rosa Li[1], Nate Kornell[2], Alexandre Pouget[1,3], and Jessica F. Cantlon[1]
[1]Department of Brain and Cognitive Sciences, University of Rochester; [2]Department of Psychology, Williams College; and [3]Department of Neuroscience, University of Geneva

Abstract
Metacognition, the ability to assess one's own knowledge, has been targeted as a critical learning mechanism in mathematics education. Yet the early childhood origins of metacognition have proven difficult to study. Using a novel nonverbal task and a comprehensive set of metacognitive measures, we provided the strongest evidence to date that young children are metacognitive. We showed that children as young as 5 years made metacognitive "bets" on their numerical discriminations in a wagering task. However, contrary to previous reports from adults, our results showed that children's metacognition is domain specific: Their metacognition in the numerical domain was unrelated to their metacognition in another domain (emotion discrimination). Moreover, children's metacognitive ability in only the numerical domain predicted their school-based mathematics knowledge. The data provide novel evidence that metacognition is a fundamental, domain-dependent cognitive ability in children. The findings have implications for theories of uncertainty and reveal new avenues for training metacognition in children.

Keywords
cognition, metacognition, childhood development, cognitive development, mathematical ability, number comprehension

Received 6/21/13; Revision accepted 5/8/14

Metacognition has been targeted as an important learning mechanism in science, technology, engineering, and particularly mathematics disciplines (Kuhn, 2000; Schoenfeld, 1992). Education research has shown that children with poor metacognitive abilities tend to overestimate their knowledge, study less, and consequently learn less than children with good metacognitive abilities (Dunlosky & Rawson, 2011; Metcalfe & Kornell, 2007). Creating interventions that explicitly teach metacognitive strategies is an active area of investigation (de Bruin & van Gog, 2012; Son, 2010). Moreover, models of mathematics instruction that include interventions on metacognitive monitoring and self-regulation, compared with models that do not, yield greater improvements on children's mathematical competence in the classroom (Garofalo & Lester, 1985; Schoenfeld, 2006).

However, it is currently unclear how metacognitive skills emerge in development or what factors influence their developmental trajectory. The vast majority of developmental research on metacognition has been with older children between the ages of 8 and 18 years, because children younger than 7 or 8 years often fail metacognitive tasks (Flavell, 1979; Reyna, 1996). Thus it is unclear what types of interventions would be effective for training metacognition to improve young children's learning within the domain of mathematics and beyond.

A long history of metacognition research in adults, especially with memory tasks, has shown that humans can access and rate their internal uncertainty (Nelson & Narens, 1990). In a typical metacognition paradigm,

Corresponding Author:
Jessica F. Cantlon, Department of Brain and Cognitive Sciences, University of Rochester, Box 270268, Meliora Hall, Rochester, NY 14620
E-mail: jcantlon@bcs.rochester.edu

subjects study a list of words or facts and estimate the likelihood that they will recall individual list items at a later time. Adults typically perform above chance when predicting their future recall of the list items, which shows that they are capable of assessing the strength of their own knowledge (Dunlosky & Bjork, 2008). An open question is to what extent metacognition is available to young children (< 8 years old). The existing data from older children and adolescents show that uncertainty monitoring improves with age (Koriat & Shitzer-Reichert, 2002; Lyons & Ghetti, 2010; Sussan & Son, 2007). However, data on the development of metacognitive skills during early childhood are generally sparse and sometimes conflicting (Balcomb & Gerken, 2008; Schneider, 2008). Several studies have relied on measures used to assess metacognition in adults, such as verbal measures of metacognitive strategies, which could be ineffective for measuring metacognitive ability in young children (Garner & Alexander, 1989; Reyna, 1996).

In the current study, we began to fill the void in the study of early childhood metacognition using a novel nonverbal metacognition task. We also investigated whether the relationship between metacognitive skill and learning in young children is content specific. Studies of adult metacognition, in which subjects rate their own knowledge across different domains, support a domain-general model of metacognitive processes (Schraw, Dunkle, Bendixen, & Roedel, 1995). A domain-general model of metacognition predicts that an individual with poor uncertainty monitoring for one domain (e.g., solving subtraction problems) will also have poor uncertainty monitoring for another domain (e.g., recognizing faces). Such cross-domain correlations in metacognitive sensitivity have been observed in adults. For example, adults who are good at estimating their knowledge of major American cities also tend to be good at estimating their knowledge of mathematical probabilities (Schraw et al., 1995; but see Kelemen, Frost, & Weaver, 2000). Additionally, adults do not show a metacognitive benefit for domains in which they have a high level of knowledge or expertise, such as music or physics, which suggests that metacognitive ability does not covary with domain knowledge (Glenberg & Epstein, 1987). Prior data indicate that adult metacognition is a general skill that is correlated across content domains and is not bound to domain knowledge.

If metacognition is also domain general in young children, this would imply that metacognitive interventions in any domain (e.g., math, reading, science) will improve metacognitive ability across all domains. Alternatively, some researchers have suggested that metacognitive abilities could be domain specific early in development and generalize across domains only as children mature (Lyons & Ghetti, 2010; Pressley, Borkowski, & Schneider, 1987; Schraw et al., 1995; Veenman & Spaans, 2005). We tested these hypotheses by comparing the metacognitive abilities of 5- to 8-year-olds in two distinct cognitive-judgment types: numerical judgments and emotional-valence judgments. We then investigated the relationship between children's metacognitive abilities and standardized measures of mathematical learning and general intelligence.

Beyond revealing the origins of metacognition, tests of metacognition in children have implications for theories of the representation of uncertainty. Recent theories of cognitive and neural representations propose that internal uncertainty is inherently encoded in perceptual representations (Knill & Pouget, 2004; Pouget, Beck, Ma, & Latham, 2013). These theories suggest that by encoding information probabilistically, the brain automatically represents both the intensity of a stimulus along a perceptual dimension and the uncertainty associated with that internal estimate of stimulus intensity (Ma, Beck, Latham, & Pouget, 2006). There is evidence that the cognitive and neural computations underlying confidence judgments, such as those tested in the current study, are derived from probabilistic representations of perceptual variables (Beck et al., 2008; Kiani & Shadlen, 2009). It is unknown whether young children represent their uncertainty during perceptual discriminations and use uncertainty to guide their postdecision confidence judgments. Evidence that young children are capable of accurately judging both the perceptual intensity of a stimulus and their uncertainty about that judgment would be consistent with a central claim from theories of probabilistic representation—that representations of uncertainty are fundamental.

Method

We asked 5- to 8-year-olds to make a basic numerosity discrimination ("Which set is larger?"), immediately followed by a retrospective wager on the accuracy of that judgment ("How sure are you?"). Children earned or lost virtual tokens depending on both their accuracy and the value of their bet. In the same session, children also made confidence judgments after perceptually comparing the valence of two facial expressions. This allowed us to directly compare children's metacognitive abilities in the number and emotion domains while keeping task demands constant. We administered a risk assessment separately to control for individual biases that might influence the children's confidence wagers. Finally, we administered standardized IQ tests to examine the relationship between metacognition and mathematics development (cf. Kelemen, Winningham, & Weaver, 2007).

Participants

We aimed to recruit 45 to 50 participants (ages 5 to 8 years), distributed approximately evenly across three age

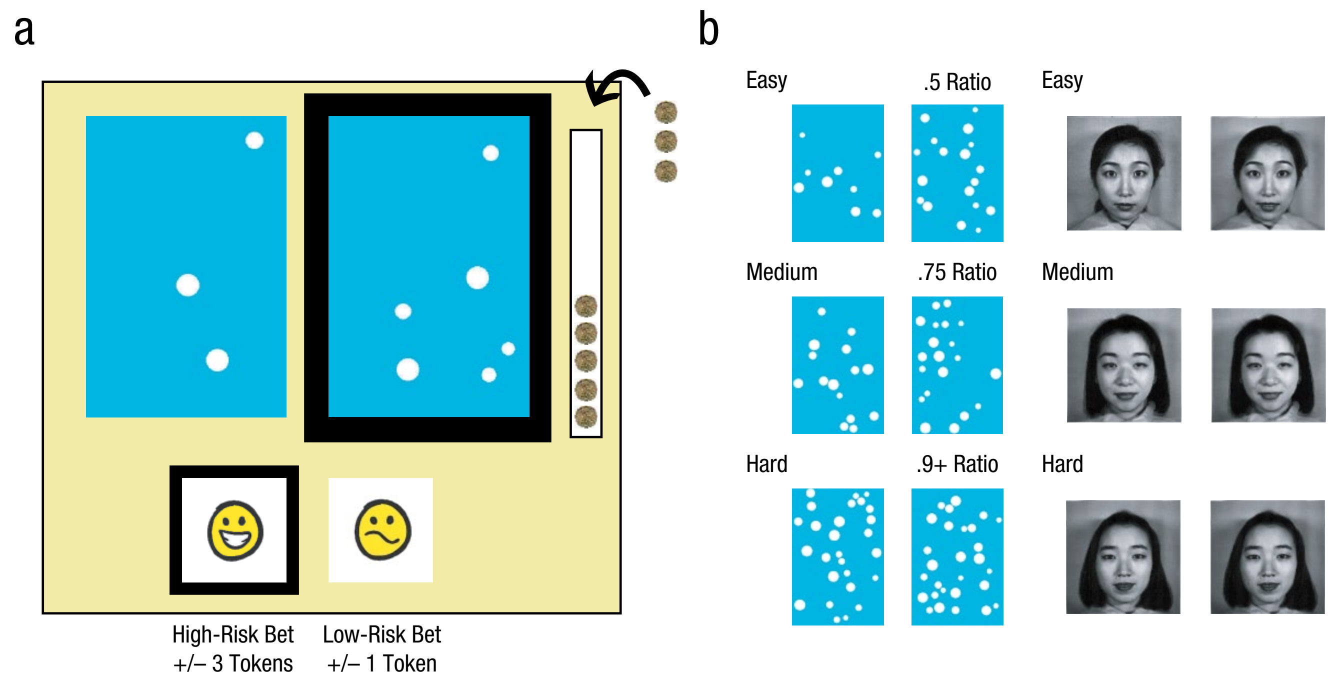

Fig. 1. Example trial in the numbers task (a) and examples of the three difficulty levels in the number and emotion tasks (b). In the numbers task, children were asked to judge which of two sets of dots had more members. Children then made a confidence judgment using one of two bet icons, a happy face or an uncertain face, which indicated a high or low bet, respectively. Both their number decision and their confidence judgment were highlighted by a black border appearing around their selection. An on-screen token counter kept track of children's wins and losses. The emotion task (not pictured here) was set up in exactly the same fashion, except that children had to judge which of two expressions from the same individual was happier. Trials in each task were classified as easy, medium, or hard on the basis of the difference in the number of dots or the closeness of the facial expressions in each pair of stimuli.

groups. We did not include 3 participants who exclusively chose one bet option during the entire task because we could not calculate their metacognitive scores. The general assessment was that these subjects were not motivated to participate or follow task instructions. Data for the remaining 45 children (mean age = 6 years, 7 months) are reported here (5- to 6-year-olds: n = 18, mean age = 5 years, 6 months, range = 5 years to 6 years, 1 month; 6- to 7-year-olds: n = 13, mean age = 6 years, 6 months, range = 6 years, 1 month to 7 years, 1 month; 7- to 8-year-olds: n = 14, mean age = 7 years, 11 months, range = 7 years, 1 month to 8 years, 11 months).

Six children exclusively chose the high bet for either the number task or the emotion task. Because some metacognitive measures are invalid for that behavior, those 6 subjects could not be included in a subset of the statistics that required a full data set for each subject (e.g., paired t tests).

Behavioral measures

Children completed a baseline risk-preference assessment, the metacognitive wagering task for both stimulus types, and finally, standardized intelligence tests. Children were rewarded with tokens in the two tasks, which they exchanged for prizes. Although they were told that their prizes would be commensurate with the amount they earned in the wagering tasks, all children received similarly valued prizes.

At the beginning of the study, children completed a version of the cups task, which we used to calculate their baseline risk preferences (Levin, Weller, Pederson, & Harshman, 2007; see the Supplemental Material available online). Next, prior to testing, children were familiarized with the metacognitive wagering paradigm (see the Supplemental Material). They then completed the two tasks, which required them to make a binary discrimination ("Which picture has more dots?" or "Which person is happier?") on a touch screen. Immediately following their decision, children made a confidence judgment by placing a wager on their accuracy (Fig. 1). Children completed at least 30 trials in each of the number and emotion tasks.

In the number task, two sets of dots were presented side by side. Within each set, the dots were randomly placed and heterogeneous in size, and on each trial, they ranged in quantity from 3 to 31. Each pair of dot arrays was classified as an easy, medium, or hard judgment based on the numerical ratio of the pair (1:2, 4:5, and 9:10 or greater, respectively). The first 10 trials varied randomly in difficulty. If children's accuracy was greater than

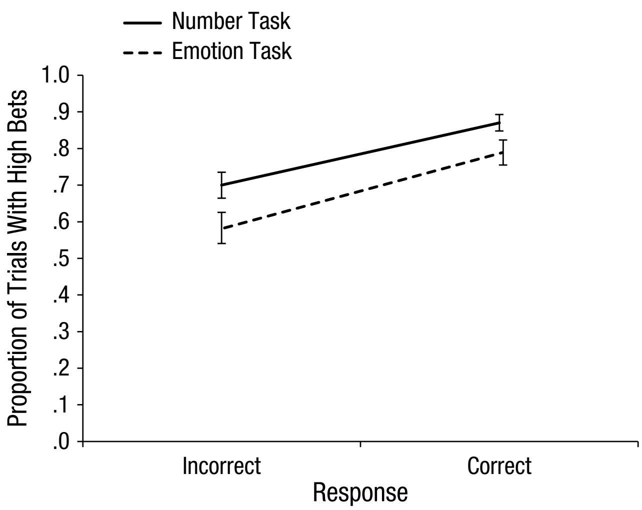

Fig. 2. Mean proportion of high bets as a function of whether children responded incorrectly or correctly in the number and emotion tasks. Error bars show standard errors of the mean.

80% on the numerical judgments in these 10 trials, perceptual-discrimination difficulty was increased to include more medium or hard judgments so that children were motivated to use both high and low bets throughout the session.

In the emotional-valence (control) task, two pictures of a single individual with different intensities in her emotional expression were presented on each trial. To systematically vary the intensity of the expression, we morphed a photograph of a neutral expression with a happy expression from the same individual in 120 steps using Morph Age Express (Version 4.1.1, Creaceed, Mons, Belgium; cf. Kelly & Metcalfe, 2011). Photographs of three female faces were taken from the Yale Face Database (Belhumeur, Hespanha, & Kriegman, 1997). Easy, medium, and hard trials were classified based on the distance between the two morphs (average distance of 48, 16, and 8 morphs, respectively). Discrimination difficulty was calibrated for each subject in the same manner as for the number task.

We calculated three measures of metacognition for each child: φ, γ, and the estimated area under the receiver-operating-characteristic (ROC) curve, or A'_{ROC} (described in detail in the Results section). The data were inspected for outliers (> 2 *SD* from the mean). One data point (a measure of A'_{ROC} from the number task) for 1 participant was the sole outlier and was excluded.

Children also completed the third edition of the Test of Early Mathematical Ability (TEMA-3; Ginsburg & Baroody, 2003) and the second edition of the Kaufman Brief Intelligence Test (KBIT-2; Kaufman & Kaufman, 2004), so we could measure their mathematical and general IQ.

Results

First, we tested children's overall metacognitive sensitivity. Then, we tested the domain specificity of children's metacognition using multiple measures of metacognitive sensitivity (φ, γ, and A'_{ROC}). Finally, we examined the correlation between children's metacognitive sensitivity and educational achievement.

Metacognitive sensitivity

Children placed appropriate bets on the accuracy of their discrimination judgments. Discrimination accuracy (as indexed by the proportion of correct responses) was significantly above chance on both the number task (M = .80), $t(44) = 21.50$, $p < .001$, and the emotion task (M = .77), $t(44) = 24.63$, $p < .001$. We found that children generally bet appropriately: They placed high-risk bets more often on trials on which they responded correctly than on trials on which they responded incorrectly, and this was true in both the number task, $t(88) = 4.08$, $p < .001$, and the emotion task, $t(88) = 3.76$, $p < .001$ (Fig. 2). In addition, the proportion of high-risk bets decreased as difficulty increased, as evidenced by a main effect of difficulty across the easy, medium, and hard conditions in the two tasks, $F(2, 42) = 30.30$, $p < .001$ (Fig. 3). A main effect of difficulty across conditions and tasks also emerged for discrimination accuracy, $F(2, 42) = 274.51$, $p < .001$.

To measure metacognitive sensitivity, we calculated the φ correlation between task accuracy (correct or incorrect responses) and confidence judgments (high or low bets) for each child (cf. Kornell, Son, & Terrace, 2007). Phi is also reducible to Pearson's r, and φ correlations significantly greater than zero reflect a pattern of successful predictions about the accuracy of one's judgments. Phi coefficients were significantly greater than zero in both the number and emotion tasks for all age groups (Table 1).

To assess effects of age, we conducted a repeated measures two-way analysis of variance on φ with age group and task as factors. Prior research with older children has reported that metacognitive sensitivity increases with age. We found a main effect of age group, $F(2, 36) = 9.73$, $p < .001$, but no effect of task, $F(1, 36) = 0.02$, $p = .90$, and no interaction, $F(2, 36) = 0.32$, $p = .73$. Metacognitive sensitivity in both the number and emotion tasks generally increased with age among 5- to 8-year-olds.

Children's metacognitive sensitivity was unaffected by their baseline risk preferences and response time cues. Children's baseline risk preferences were not correlated with φ for either the number or the emotion task (Table S2 in the Supplemental Material). This indicates that

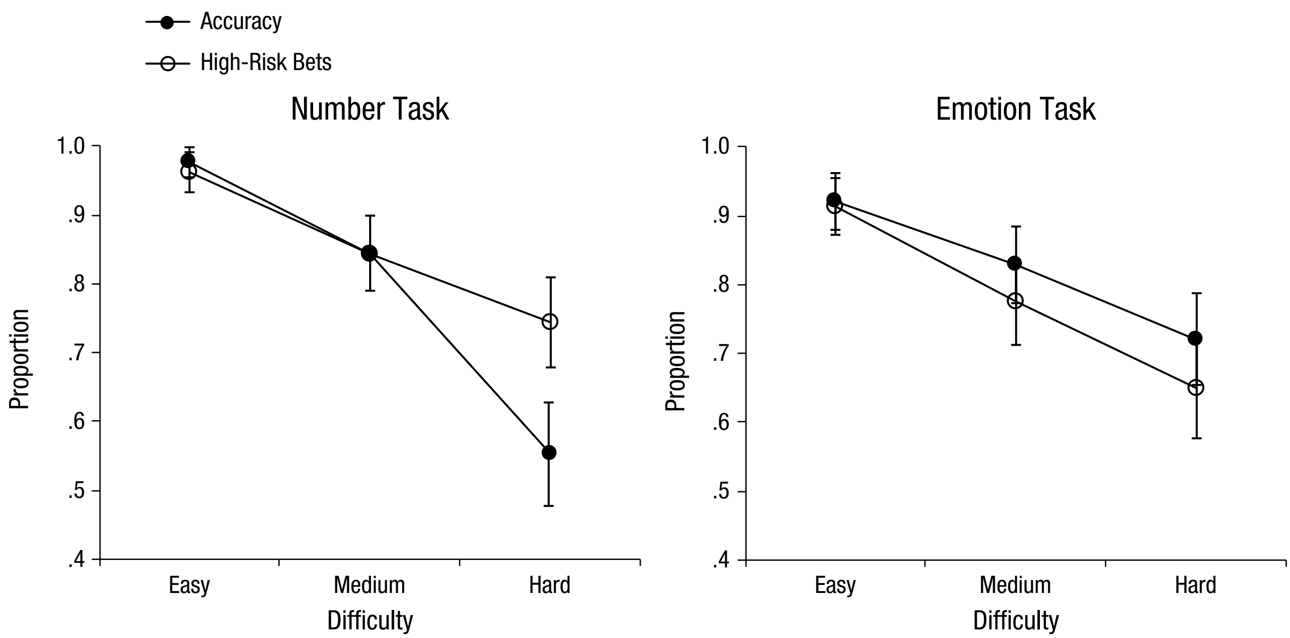

Fig. 3. Mean proportion of accurate responses and high-risk bets as a function of trial difficulty, separately for the number and emotion tasks. Error bars show standard errors of the mean.

children's metacognitive judgments were not dominated by generic risk-seeking or risk-aversive behaviors.

According to some models of decision making, the elapsed time to make a perceptual judgment (response time) is a critical internal parameter for estimating uncertainty (Kiani & Shadlen, 2009). However, other models have posited that confidence judgments can be confounded by response time in wagering tasks when the appropriate bet can be predicted by the "public cue" of one's own response time (Hampton, 2009; Koriat & Ackerman, 2010). That is, subjects could base their confidence judgments on the external observation of their own response time, rather than by using an internal monitoring mechanism.

To test whether response time was a critical predictor of children's confidence judgments, we calculated φ as a partial correlation of accuracy and risk that controlled for discrimination response time. The φ correlation remained significant for both the number task (M = .21), $t(40)$ = 6.26, p < .001, and the emotion task (M = .17), $t(42)$ = 4.80, p < .001, which suggests that children's confidence judgments depended on an internal representation of uncertainty that is independent of response time (as predicted in uncertainty models developed by Beck et al., 2008).

Table 1. Tests of Metacognitive Sensitivity by Age Group and Task

	φ			γ			A'_{ROC}		
Task and age group	M	t	p	M	t	p	M	t	p
Number									
5- to 6-year-olds	.13	$t(16)$ = 2.33	.03	.20	$t(16)$ = 1.23	.24	.64	$t(13)$ = 2.93	.01
6- to 7-year-olds	.22	$t(10)$ = 4.82	< .001	.52	$t(10)$ = 4.67	< .001	.68	$t(10)$ = 4.32	< .001
7- to 8-year-olds	.35	$t(12)$ = 7.55	< .001	.74	$t(12)$ = 13.75	< .001	.78	$t(12)$ = 14.82	< .001
All participants	.22	$t(40)$ = 6.84	< .001	.46	$t(40)$ = 5.55	< .001	.68	$t(38)$ = 6.36	< .001
Emotion									
5- to 6-year-olds	.12	$t(16)$ = 2.83	.01	.26	$t(16)$ = 1.53	.15	.68	$t(13)$ = 5.41	< .001
6- to 7-year-olds	.29	$t(11)$ = 3.26	.008	.41	$t(11)$ = 1.86	.09	.75	$t(9)$ = 5.22	< .001
7- to 8-year-olds	.30	$t(13)$ = 7.39	< .001	.66	$t(13)$ = 14.39	< .001	.77	$t(13)$ = 13.32	< .001
All participants	.22	$t(42)$ = 6.53	< .001	.43	$t(42)$ = 4.62	< .001	.73	$t(37)$ = 11.75	< .001

Note: All t tests are two-tailed comparisons against chance—0 for φ and γ; 0.5 for the area under the receiver-operating-characteristic (ROC) curve, or A'_{ROC}.

Table 2. Correlations of Each of the Three Measures of Metacognitive Sensitivity Between the Number and Emotion Tasks

Measure	*r*	*p*
φ	$r(37) = -.12$	.91
γ	$r(37) = .08$	.64
A'_{ROC}	$r(30) = .04$	.84

Note: A'_{ROC} = area under the receiver-operating-characteristic (ROC) curve.

Domain-specificity of metacognition

To test whether metacognitive knowledge follows a domain-general or domain-specific trajectory during early childhood, we compared several measures of metacognitive sensitivity across the number and emotion domains: confidence bias, φ, γ, and A'_{ROC}. Prior studies of metacognition have used only one or two of these measures (Fleming & Dolan, 2012; Masson & Rotello, 2009; Nelson, 1984). From a statistical perspective, each measure has unique strengths and weaknesses. Here, we combined all three measures to provide a robust test of metacognitive ability in children.

We calculated a measure of confidence-judgment bias by subtracting each child's average task accuracy from his or her average confidence judgment (high or low) (number task: $M = .05$, $SD = .046$; emotion task: $M = -.08$, $SD = .076$). Biases differed significantly between the number and emotion domains, $t(38) = -3.45$, $p = .001$, with children exhibiting marginal overconfidence in their numerical judgments and underconfidence in their emotion judgments—one-sample *t* tests: $t(38) = 1.92$, $p = .059$; $t(38) = -2.10$, $p = .04$; see the Supplemental Material for further tests. However, bias effects are known to be influenced by differences in task accuracy, so the true degree of children's overestimation or underestimation is unclear (Schraw & Roedel, 1994).

The φ coefficient, as described earlier, represents the correlation between each subject's discrimination accuracy and his or her risk choices. Children's φ values in the number and emotion domains were not correlated, which suggests that children's metacognitive sensitivity is not uniform across different judgment types (Table 2; see the Supplemental Material for further discussion).

The γ coefficient is a nonparametric measure of correlation that is calculated by taking the difference between concordances (e.g., high bets on correctly identified items) and discordances (e.g., high bets on incorrectly identified items) and dividing by the total number of trials, bounding the score between 1 and −1 (Nelson, 1984). The γ coefficient was significantly above zero for both tasks and for most age groups, although the youngest children showed large variance in their scores (Table 1).

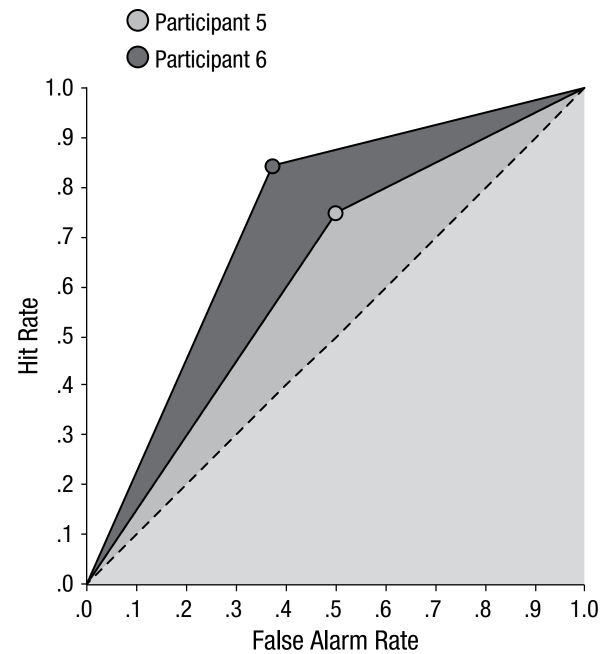

Fig. 4. Example receiver-operating-characteristic (ROC) functions using data from 2 participants in the emotion task. The ROC estimate shows the mean proportion of hits as a function of the mean proportion of false alarms. The dashed line represents metacognitive insensitivity (i.e., the hit rate equals the false alarm rate). Any deviations toward the upper left of the graph indicate sensitivity above chance.

However, the lack of a γ effect in young children should be viewed cautiously because of systematic biases in the γ measure (Masson & Rotello, 2009). As with φ, γ was not statistically predicted by children's baseline risk preference (Table S2). Gamma also was not correlated between the number and emotion domains (Table 2).

We also calculated a nonparametric measure of metacognitive sensitivity from signal detection theory, A'_{ROC} (Galvin, Podd, Drga, & Whitmore, 2003; Kornbrot, 2006). The ROC function plots the hit rate (high bets on trials with correct responses or low bets on trials with incorrect responses, or *concordances*) against the false alarm rate (low bets on trials with correct responses and high bets on trials with incorrect responses, or *discordances*), such that equivalence between the two represents no metacognitive sensitivity ($A'_{ROC} = .5$; see Fig. 4).

A'_{ROC} was significantly greater than chance for both tasks and all age groups, which indicates that even the youngest children showed significant metacognitive sensitivity (Table 1; also see Table S1 in the Supplemental Material). A'_{ROC} was not significantly influenced by baseline risk preferences (Table S2). And as with φ and γ, A'_{ROC} was not correlated between the number and emotion domains (Table 2).

The results from φ, γ, and A'_{ROC} broadly indicate that individual differences in metacognitive ability are not correlated across content domains during early childhood.

Table 3. Loadings of Each Metacognition Measure on the Number and Emotion Tasks

Task and measure	Component 1 (n = 32)	Component 2 (n = 32)
Number		
φ	.968*	.026
γ	.984*	.019
A'_{ROC}	.987*	−.026
Emotion		
φ	−.153	.923*
γ	.109	.951*
A'_{ROC}	.062	.991*
Eigenvalue	2.93	2.73
Variance explained	49%	46%

Note: Loadings are reported for the varimax-rotated solution. Principal component analysis was performed on the correlation matrix to make the scales equivalent across variables. A'_{ROC} = area under the receiver-operating-characteristic (ROC) curve.
*$p < .005$.

To test whether this was the broad pattern across all measures, we performed a canonical correlation analysis (CCA; Hair, Anderson, Tatham, & Black, 1998). CCA finds the maximal correlation between a linear combination of one set of variables (in this case, number-related metacognition measures) and a linear combination of another set (in this case, emotion-related measures). The full CCA model indicated that there was no significant linear relationship between the two domains (Wilks Λ = .759), $F(9, 63.43) = 0.846$, $p = .58$. This supports the conclusion that metacognitive ability is domain specific in children.

We next used principal component analysis (PCA) to obtain a summary score of metacognitive sensitivity for each subject. PCA plots the data in an n-dimensional space (n = number of input variables) and determines which of the n dimensions, or components, accounts for most of the variance in the data (Hair et al., 1998).

We performed a PCA across the three measures of metacognitive sensitivity (φ, γ, and A'_{ROC}) for the two domains, which resulted in six input variables. Both the Kaiser criterion and the scree plot indicated that two components accounted for most of the variance in the data (Table 3). The loads, which quantify the strength of the relationship between the component and each input variable, were varimax-rotated to improve the interpretability of the solution. The metacognitive measures clustered together by domain. All number measures loaded highly on Component 1, whereas all emotion measures loaded highly on Component 2 (Table 3). These results are consistent with the conclusion that children's performance segregates by content domain, rather than by measurement type. We then used the individual PCA scores for each component as a summary measure of number-related or emotion-related metacognition to test the relationship between metacognitive sensitivity and children's cognitive development.

Metacognition and education

To investigate the possible link between domain-specific metacognition and cognitive development, we correlated children's metacognitive scores from the PCA with their school-based test scores. Metacognitive sensitivity on numerical judgments, but not on emotion judgments, correlated with children's mathematics test scores on the TEMA-3—number task: $r(30) = .52$, $p < .001$; emotion task: $r(30) = .26$, $p = .17$ (Fig. 5). Because children's mathematics ability was not broadly related to metacognitive sensitivity, the effect is not likely explained by a generic age or competence effect.

No measure of metacognition was significantly correlated with general intelligence scores as measured by the KBIT-2—emotion task: $r(30) = .11$, $p = .55$; number task: $r(30) = .13$, $p = .46$. This is further evidence of the domain specificity of metacognition in that children's numerical metacognition is uniquely related to their mathematical skills (Cantlon, 2012).

Discussion

Our data provide novel evidence that (a) young children are capable of reporting their uncertainty in a nonverbal metacognitive task; (b) uncertainty monitoring in early childhood is not a global ability that matures uniformly across content domains, but instead develops along domain-specific trajectories; and (c) children's domain-specific metacognition for numerical discrimination predicts their formal education achievement in mathematics.

Young children monitor their internal uncertainty

Several prior studies have reported that young children fail metacognitive tasks (Flavell, 1979; Flavell, Green, & Eleanor, 2000; Miller & Bigi, 1989; Myers & Paris, 1978; for a review, see Garner & Alexander, 1989; Reyna, 1996). However, we showed with a nonverbal wagering task that young children are capable of uncertainty monitoring for a basic perceptual discrimination task. That is, young children were metacognitively sensitive across age groups, stimulus types, and multiple metacognitive measures (φ, γ, A'_{ROC}). This robust relationship between discrimination accuracy and confidence judgments on a trial-by-trial basis provides strong evidence that young children can access and track an internal estimate of their uncertainty.

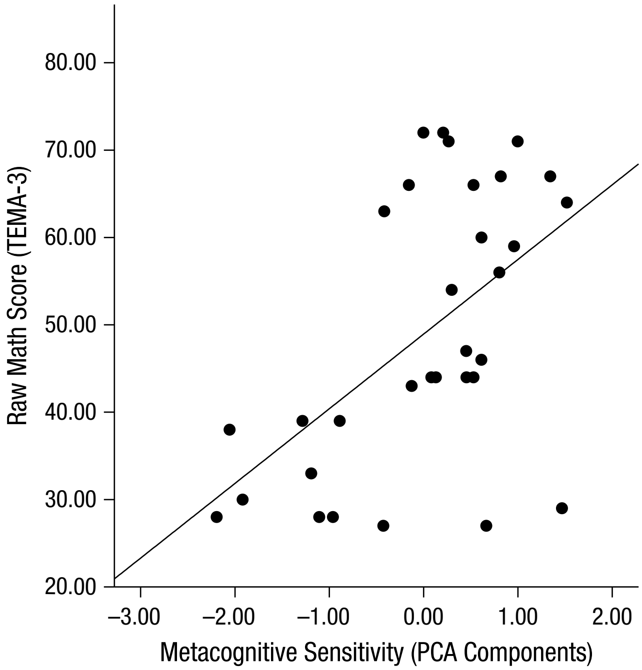

Fig. 5. Scatter plot (with best-fitting regression line) showing the correlation between math score on the third edition of the Test of Early Mathematical Ability (TEMA-3; Ginsburg & Baroody, 2003) and score from a principal component analysis (PCA) performed on three different measures of metacognition (φ, γ, and area under the receiver-operating-characteristic curve) from the number task.

Uncertainty monitoring is domain specific in early childhood

We found no correlation between children's metacognitive abilities in the number and emotion domains across multiple measures of metacognitive sensitivity. This was confirmed by CCA and PCA. This finding suggests that metacognitive skill does not globally mature across content domains.

Furthermore, we found that generic metacognitive skill cannot predict formal mathematical ability—only numerical metacognition, not emotion-related metacognition, correlated with scores on a standardized math test. Whether numerical metacognition drives the development of mathematical ability or whether mathematical ability drives the development of numerical metacognition remains to be determined. Nonetheless, these results imply that metacognition develops along domain-specific trajectories, such that children's metacognitive abilities depend on information content and domain knowledge.

Our results support the hypothesis that metacognition transitions from a domain-specific process to a domain-general mechanism over development. A previous study with older children suggests that metacognitive abilities begin to show correlations among different content domains during adolescence (Veenman & Spaans, 2005). By adulthood, metacognition seems to be largely domain general and independent of domain knowledge (Glenberg & Epstein, 1987; Schraw et al., 1995). These data are broadly consistent with the theory that content-specific learning strategies (and therefore, metacognitive abilities) develop as knowledge increases in a domain (Pressley et al., 1987; Reyna, 1996). Global abilities emerge when children can generalize knowledge structures and learning procedures across domains (Lyons & Ghetti, 2010; Schneider, 2008).

Implications for education

Our study shows that children can report their uncertainty nonverbally by at least the age of 5 years. This suggests that nonverbal paradigms, such as wagering, could be used in metacognitive interventions with preschool children. Yet although children generally showed metacognition in our task, their metacognitive judgments were still imperfect. Children made several errors in their metacognitive bets, and on average, these errors were biased toward higher wagers (overconfidence) for numerical judgments and lower wagers for emotion judgments (underconfidence; see the Supplemental Material for discussion). Moreover, as described in the Results section, children's metacognitive skill was domain specific and related to their formal domain knowledge. Overall, our data suggest that young children can show different metacognitive skills and biases for different stimulus types. In this case, generic interventions to improve children's metacognitive accuracy might be less effective than training specific metacognitive strategies that depend on the content of the learning materials.

We predict that early in development, the types of metacognitive errors that children make in one domain will not transfer to another domain. Characterizing these errors could help educators develop specific intervention techniques (Borkowski, 1992; Carr & Biddlecomb, 1998; Kuhn, 2000).

Children's internal uncertainty may be represented probabilistically

Questions remain as to the precise nature of the representations that allow children to monitor their uncertainty. There are currently few formal models of metacognitive processes for children or adults (Koriat & Ackerman, 2010; Nelson & Narens, 1990). One possibility, suggested by probabilistic models of representation, is that humans inherently encode uncertainty in their perceptual representations (Knill & Pouget, 2004; Pouget et al., 2013). If this is true, then even young children should show evidence of representing the uncertainty associated with their perceptual judgments. Our data show that they do.

The probabilistic theory of representation proposes that the representation of uncertainty is a fundamental component of the cognitive architecture (Beck et al., 2008; Kiani & Shadlen, 2009; Ma et al., 2006; Pouget et al., 2013). Our finding that young children accurately report their uncertainty is consistent with the proposal that the representation of uncertainty is a fundamental ability. Developmental changes between childhood and adulthood have not yet been studied with this framework. Yet probabilistic theories could provide new insights into the nature of fundamental mechanisms by which uncertainty is represented and those by which uncertainty is consciously accessed and reported.

Conclusion

We conclude that young children "know what they know" in making basic perceptual judgments. Young children can accurately report their uncertainty nonverbally, by at least the age of 5 years. Yet children's metacognitive abilities continue to develop into childhood and adolescence, and even adults are far from perfect (Dunlosky & Bjork, 2008). Our data suggest that metacognitive abilities develop in tandem with domain-specific changes in children's knowledge. The links we observed between children's metacognitive sensitivity and formal mathematical abilities indicate that early interventions on children's metacognitive strategies could have far-reaching effects on their education.

Author Contributions

All authors contributed to the study concept and design. Testing and data collection were performed by V. A. Vo and R. Li. Data analysis and interpretation were done by V. A. Vo and R. Li under the supervision of J. F. Cantlon, A. Pouget, and N. Kornell. All authors approved the final version of the manuscript before submission.

Acknowledgments

We are grateful to Brad Mahon, Jeff Beck, the Aslin Lab, and the members of the Concepts, Actions, and Objects Lab, especially Gina Gerhardt, Laura Ackerman, and Sydney Robinson.

Declaration of Conflicting Interests

The authors declared that they had no conflicts of interest with respect to their authorship or the publication of this article.

Funding

This work was funded by a grant from the National Science Foundation Research and Evaluation on Education in Science and Engineering program (1109366) to A. Pouget and J. F. Cantlon and by a grant from the National Institutes of Health (R01 HD064636) to J. F. Cantlon.

Supplemental Material

Additional supporting information may be found at http://pss.sagepub.com/content/by/supplemental-data

References

Balcomb, F. K., & Gerken, L. (2008). Three-year-old children can access their own memory to guide responses on a visual matching task. *Developmental Science*, *11*, 750–760. doi:10.1111/j.1467-7687.2008.00725.x

Beck, J. M., Ma, W. J., Kiani, R., Hanks, T., Churchland, A. K., Roitman, J., . . . Pouget, A. (2008). Probabilistic population codes for Bayesian decision making. *Neuron*, *60*, 1142–1152. doi:10.1016/j.neuron.2008.09.021

Belhumeur, P. N., Hespanha, J. P., & Kriegman, D. J. (1997). Eigenfaces vs. Fisherfaces: Recognition using class specific linear projection. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, *19*, 711–720.

Borkowski, J. G. (1992). Metacognitive theory: A framework for teaching literacy, writing, and math skills. *Journal of Learning Disabilities*, *25*, 253–257. doi:10.1177/002221949202500406

Cantlon, J. F. (2012). Math, monkeys, and the developing brain. *Proceedings of the National Academy of Sciences, USA*, *109*(Suppl. 1), 10725–10732.

Carr, M., & Biddlecomb, B. (1998). Metacognition in mathematics from a constructivist perspective. In D. J. Hacker, J. Dunlosky, & A. C. Graesser (Eds.), *Metacognition in educational theory and practice* (pp. 69–91). Mahwah, NJ: Erlbaum.

de Bruin, A. B. H., & van Gog, T. (2012). Improving self-monitoring and self-regulation: From cognitive psychology to the classroom. *Learning and Instruction*, *22*, 245–252. doi:10.1016/j.learninstruc.2012.01.003

Dunlosky, J., & Bjork, R. A. (2008). The integrated nature of metamemory and memory. In J. Dunlosky & R. A. Bjork (Eds.), *Handbook of metamemory and memory* (pp. 11–28). Hillsdale, NJ: Psychology Press.

Dunlosky, J., & Rawson, K. A. (2011). Overconfidence produces underachievement: Inaccurate self evaluations undermine students' learning and retention. *Learning and Instruction*, *22*, 271–280. doi:10.1016/j.learninstruc.2011.08.003

Flavell, J. H. (1979). Metacognition and cognitive monitoring: A new area of cognitive–developmental inquiry. *American Psychologist*, *34*, 906–911.

Flavell, J. H., Green, F. L., & Eleanor, R. (2000). Development of children's awareness of their own thoughts. *Journal of Cognition and Development*, *1*, 97–112. doi:10.1207/S15327647JCD0101N

Fleming, S. M., & Dolan, R. J. (2012). The neural basis of metacognitive ability. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *367*, 1338–1349. doi:10.1098/rstb.2011.0417

Galvin, S. J., Podd, J. V., Drga, V., & Whitmore, J. (2003). Type 2 tasks in the theory of signal detectability: Discrimination between correct and incorrect decisions. *Psychonomic Bulletin & Review*, *10*, 843–876.

Garner, R., & Alexander, P. A. (1989). Metacognition: Answered and unanswered questions. *Educational Psychologist*, *24*, 143–158.

Garofalo, J., & Lester, F. K. (1985). Metacognition, cognitive monitoring, and mathematical performance. *Journal for Research in Mathematics Education*, *16*, 163–176.

Ginsburg, H. P., & Baroody, A. J. (2003). *Test of Early Mathematical Ability* (3rd ed.). Austin, TX: Pro-Ed.

Glenberg, A. M., & Epstein, W. (1987). Inexpert calibration of comprehension. *Memory & Cognition*, *15*, 84–93.

Hair, J. F., Anderson, R. E., Tatham, R. L., & Black, W. C. (1998). *Multivariate data analysis* (5th ed.). Upper Saddle River, NJ: Pearson.

Hampton, R. R. (2009). Multiple demonstrations of metacognition in nonhumans: Converging evidence or multiple mechanisms? *Comparative Cognition & Behavior Reviews*, *4*, 17–28. doi:10.3819/ccbr.2009.40002

Kaufman, A. S., & Kaufman, N. L. (2004). *Kaufman Brief Intelligence Test* (2nd ed.). Upper Saddle River, NJ: Pearson.

Kelemen, W. L., Frost, P. J., & Weaver, C. A. (2000). Individual differences in metacognition: Evidence against a general metacognitive ability. *Memory & Cognition*, *28*, 92–107.

Kelemen, W. L., Winningham, R. G., & Weaver, C. A. (2007). Repeated testing sessions and scholastic aptitude in college students' metacognitive accuracy. *European Journal of Cognitive Psychology*, *19*, 689–717.

Kelly, K. J., & Metcalfe, J. (2011). Metacognition of emotional face recognition. *Emotion*, *11*, 896–906. doi:10.1037/a0023746

Kiani, R., & Shadlen, M. N. (2009). Representation of confidence associated with a decision by neurons in the parietal cortex. *Science*, *324*, 759–764. doi:10.1126/science.1169405

Knill, D. C., & Pouget, A. (2004). The Bayesian brain: The role of uncertainty in neural coding and computation. *Trends in Neurosciences*, *27*, 712–719. doi:10.1016/j.tins.2004.10.007

Koriat, A., & Ackerman, R. (2010). Choice latency as a cue for children's subjective confidence in the correctness of their answers. *Developmental Science*, *13*, 441–453. doi:10.1111/j.1467-7687.2009.00907.x

Koriat, A., & Shitzer-Reichert, R. (2002). Metacognitive judgments and their accuracy: Insights from the processes underlying judgments of learning in children. In P. Chambres, M. Izaute, & P.-J. Marescaux (Eds.), *Metacognition: Processes, function, and use* (pp. 1–17). Norwell, MA: Kluwer.

Kornbrot, D. E. (2006). Signal detection theory, the approach of choice: Model-based and distribution-free measures and evaluation. *Perception & Psychophysics*, *68*, 393–414.

Kornell, N., Son, L. K., & Terrace, H. S. (2007). Transfer of metacognitive skills and hint seeking in monkeys. *Psychological Science*, *18*, 64–71. doi:10.1111/j.1467-9280.2007.01850.x

Kuhn, D. (2000). Metacognitive development. *Current Directions in Psychological Science*, *9*, 178–181.

Levin, I. P., Weller, J. A., Pederson, A. A., & Harshman, L. A. (2007). Age-related differences in adaptive decision making: Sensitivity to expected value in risky choice. *Judgment and Decision Making*, *2*, 225–233.

Lyons, K. E., & Ghetti, S. (2010). Metacognitive development in early childhood: New questions about old assumptions. In A. Efklides & P. Misailidi (Eds.), *Trends and prospects in metacognition research* (pp. 259–278). Boston, MA: Springer. doi:10.1007/978-1-4419-6546-2

Ma, W. J., Beck, J. M., Latham, P. E., & Pouget, A. (2006). Bayesian inference with probabilistic population codes. *Nature Neuroscience*, *9*, 1432–1438. doi:10.1038/nn1790

Masson, M. E. J., & Rotello, C. M. (2009). Sources of bias in the Goodman-Kruskal gamma coefficient measure of association: Implications for studies of metacognitive processes. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *35*, 509–527. doi:10.1037/a0014876

Metcalfe, J., & Kornell, N. (2007). Principles of cognitive science in education: The effects of generation, errors, and feedback. *Psychonomic Bulletin & Review*, *14*, 225–229.

Miller, P. H., & Bigi, L. (1989). The development of children's understanding of attention. *Merrill-Palmer Quarterly of Behavior and Development*, *25*, 235–250.

Myers, M. I., & Paris, S. G. (1978). Children's metacognitive knowledge about reading. *Journal of Educational Psychology*, *70*, 680–690.

Nelson, T. O. (1984). A comparison of current measures of the accuracy of feeling-of-knowing predictions. *Psychological Bulletin*, *95*, 109–133.

Nelson, T. O., & Narens, L. (1990). Metamemory: A theoretical framework and new findings. In G. H. Bower (Ed.), *The psychology of learning and motivation* (Vol. 26, pp. 125–173). San Diego, CA: Academic Press.

Pouget, A., Beck, J. M., Ma, W. J., & Latham, P. E. (2013). Probabilistic brains: Knowns and unknowns. *Nature Neuroscience*, *16*, 1170–1178. doi:10.1038/nn.3495

Pressley, M., Borkowski, J. G., & Schneider, W. (1987). Cognitive strategies: Good strategy users coordinate metacognition and knowledge. In R. Vasta (Ed.), *Annals of child development* (Vol. 4, pp. 89–129). Greenwich, CT: JAI Press.

Reyna, V. F. (1996). Conceptions of memory development with implications for reasoning and decision making. In R. Vasta (Ed.), *Annals of child development* (Vol. 12, pp. 87–118). Greenwich, CT: JAI Press.

Schneider, W. (2008). The development of metacognitive knowledge in children and adolescents: Major trends and implications for education. *Mind, Brain, and Education*, *2*, 114–121. doi:10.1111/j.1751-228X.2008.00041.x

Schoenfeld, A. H. (1992). Learning to think mathematically: Problem solving, metacognition, and sense-making in mathematics. In D. A. Grouws (Ed.), *Handbook for research on mathematics teaching and learning* (pp. 334–370). New York, NY: Macmillan.

Schoenfeld, A. H. (2006). Mathematics teaching and learning. In P. A. Alexander & P. H. Winne (Eds.), *Handbook of educational psychology* (pp. 479–510). New York, NY: Routledge.

Schraw, G., Dunkle, M. E., Bendixen, L. D., & Roedel, T. D. (1995). Does a general monitoring skill exist? *Journal of Educational Psychology*, *87*, 433–444.

Schraw, G., & Roedel, T. D. (1994). Test difficulty and judgment bias. *Memory & Cognition*, *22*, 63–69.

Son, L. K. (2010). Metacognitive control and the spacing effect. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *36*, 255–262. doi:10.1037/a0017892

Sussan, D., & Son, L. K. (2007). The training of metacognitive monitoring in children. *Columbia Undergraduate Science Journal*, *2*(1), 98–109.

Veenman, M. V. J., & Spaans, M. A. (2005). Relation between intellectual and metacognitive skills: Age and task differences. *Learning and Individual Differences*, *15*, 159–176. doi:10.1016/j.lindif.2004.12.001