DOI: 10.1111/cdev.13714

EMPIRICAL ARTICLE

CHILD DEVELOPMENT

The development of uncertainty monitoring during kindergarten: Change and longitudinal relations with executive function and vocabulary in children from low-income backgrounds

Christopher R. Gonzales[1] | **Alexis Merculief**[2] | **Megan M. McClelland**[2] | **Simona Ghetti**[1]

[1]Center for Mind and Brain, University of California, Davis, Davis, California, USA
[2]Human Development and Family Sciences, Oregon State University, Corvallis, Oregon, USA

Correspondence
Christopher R. Gonzales, Center for Mind and Brain, University of California, Davis, Davis, CA, USA.
Email: chrgonzales@ucdavis.edu

Funding information
Institute of Education Sciences, Grant/Award Number: R305A150192; Eunice Kennedy Shriver National Institute of Child Health and Human Development, Grant/Award Number: 1F32HD097896-01

Abstract
Children's ability to monitor subjective feelings of uncertainty (i.e., engage in uncertainty monitoring) is a central metacognitive skill. In the current study, we examined the development of uncertainty monitoring as well as its relations with vocabulary and executive function development in children (N = 137, 52% female) from predominately White and Latinx/Hispanic backgrounds when they were 4–6 years old and enrolled in a Head Start preschool and kindergarten between 2018 and 2019. We found that children's uncertainty monitoring improved during the kindergarten year. Children's executive function and vocabulary in preschool and vocabulary growth from preschool to kindergarten predicted uncertainty monitoring at the end of kindergarten, which sheds new light on potential mechanisms supporting children's metacognitive development.

Metacognitive development has long been held as a driving force behind improvements in children's social and cognitive skills and is seen as foundational for academic achievement (Flavell, 1979; Koriat & Goldsmith, 1996). An important aspect of children's metacognitive development is the ability to accurately monitor ongoing subjective feelings of uncertainty (i.e., engage in uncertainty monitoring), which may underlie children's curiosity about the world (Dunlosky & Metcalfe, 2008; Ronfard et al., 2017) and information-seeking behaviors (Selmeczy et al., 2021). Uncertainty monitoring is readily evident by middle childhood (Dunlosky & Rawson, 2012; Fandakova et al., 2017; Lockl & Schneider, 2007; Roebers et al., 2007), but research on the development of this ability in early childhood is much scarcer. There are only a few empirical studies which have illustrated young children's ability to engage in uncertainty monitoring in simple cognitive tasks (Coughlin et al., 2015; Hembacher & Ghetti, 2014; Lyons & Ghetti, 2011, 2013), in which young children verbally reported lower certainty when they made a mistake compared to higher certainty when they made an accurate decision. The relation between uncertainty monitoring and other variables during this age period is also not well understood. In the current study, we sought to expand upon this work in at least two ways. First, we documented the development of uncertainty monitoring with a longitudinal design in a sample of young children from families with low income, which is a typically underrepresented demographic in research on metacognition. Second, we examined the relation between uncertainty monitoring and children's executive function and vocabulary, which have both proven to be important predictors in many areas of cognitive and social development during the transition to kindergarten (McClelland, Cameron, Connor, et al., 2007; Schmitt et al., 2019).

Abbreviations: CFI, comparative fit index; ELL, English Language Leaner; FIML, full information maximum likelihood; HTKS-R, Head–Toes–Knees–Shoulders Revised; MAR, missing at random; RMSEA, root mean square error of approximation; SDT, signal detection theory; SEM, structural equation modeling; WEIRD, western, educated, industrialized, rich, and democratic; WJ-III, Woodcock–Johnson-III.

Uncertainty monitoring in early childhood

Uncertainty monitoring is the ability to reflect on feelings of certainty or uncertainty associated with ongoing mental operations (Lyons & Ghetti, 2010). Uncertainty monitoring associated with one's own learning may be the basis for such decisions as seeking help from a teacher or peer and is thus central for optimal self-guided learning (Zimmerman & Schunk, 2012). Children's ability to engage in uncertainty monitoring can guide the effectiveness of these types of behaviors, especially when children face difficult test questions or are given ambiguous instructions (Koriat & Goldsmith, 1996). The link between uncertainty monitoring and the display of behaviors that might facilitate learning has also been demonstrated during the preschool years. For example, when given a simple perceptual judgment task (e.g., choosing which of two degraded line drawings is a bunny), Lyons and Ghetti (2011, 2013) found that 3- to-5-year-old children were more confident on average in trials with an accurate perceptual judgment versus an inaccurate perceptual judgment, and they were also more likely to strategically withhold answers on trials in which they were least confident to receive a better reward at the end of the task. Expanding on this finding, Coughlin and colleagues (Coughlin et al., 2015) found that preschool children were more likely to ask for help from an adult on trials where they expressed lower confidence levels versus trials where they expressed high confidence levels. Furthermore, Hembacher and Ghetti (2014) found a similar pattern of results with preschool-aged children's performance on a simple forced-choice memory task, suggesting that children's uncertainty monitoring ability and its relation to strategic learning behavior may be robust across different types of cognitive tasks. Together, these studies provide compelling evidence for the presence of uncertainty monitoring in young children, and for the potential role of this ability to set the foundation and propel young children's learning during the critical transition to formal education settings.

Prior studies have suggested age-related improvements in uncertainty monitoring during the preschool years (Hembacher & Ghetti, 2014; Lyons & Ghetti, 2011) with older children demonstrating larger mean differences in confidence ratings between accurate and inaccurate trials compared to younger children, but the examination of within-individual change over time is necessary to elucidate how this improvement occurs. Results in prior studies have shown that younger children are often overconfident in their incorrect responses (e.g., Lipko et al., 2009), suggesting that improvements in children's uncertainty monitoring may reflect children's ability to recognize when they feel less confident after making mistakes or experiencing gaps in understanding with relatively little change in confidence for correct judgments. However, it is important to note that although a high proportion of children may be at ceiling in their confidence ratings for correct answers, changes in children's uncertainty monitoring could also result from the subset of children who are not at ceiling, becoming more confident after making accurate decisions because they may learn to identify the most informative cues to accuracy. Thus, the development of uncertainty monitoring may not only reflect children's developing capacity to recognize errors and reduce overconfidence for them (Lipko et al., 2009), but may also bring about some combination of the ability to recognize uncertainty after making mistakes as well as to experience greater certainty after accurate decisions.

Although metacognition is often conceptualized as a universal and generalizable force in children's development (Flavell, 1979; Koriat & Goldsmith, 1996), studies on children's early uncertainty monitoring have examined samples of children from families with predominantly White and upper middle-class backgrounds. This is not unusual for research in both cognitive research (Roberts et al., 2020; Rowley & Camacho, 2015) and child development more broadly (Nielsen et al., 2017), but the persistence of these sample characteristics in developmental studies can potentially limit the generalizability of research findings. An overreliance on samples from WEIRD (western, educated, industrialized, rich, and democratic) populations is a recognized problem in many areas of research in the social sciences (Henrich et al., 2010), and this problem may be particularly consequential in developmental psychology where systematic differences across cultures or social classes can have major impacts on both the time course and relative patterns of development found within children. Research on the development of different cognitive skills has demonstrated that children from low-income or other types of disadvantaged families are often faced with negative factors such as racism, lack of access to necessary resources and other types of inequalities which can interfere with development (Blair & Raver, 2012; Sektnan et al., 2010). Thus, it is important to demonstrate that uncertainty monitoring in early childhood has the potential to be a robust and generalizable phenomenon in development even across children from different types of backgrounds.

Assessing level and change in uncertainty monitoring

There are several approaches to assess uncertainty monitoring ranging from simple correlations between task accuracy and confidence reports to more complex modeling approaches (Fleming & Lau, 2014). One approach in studies with young children has been to compare children's average levels of confidence for accurate compared to inaccurate trials in different types of simple cognitive tasks (Ghetti et al., 2013). This comparison provides an intuitive means to appreciate whether children's experience different degrees of certainty depending on the accuracy of cognitive acts and the source of this difference. For example, studies that found age-related increases in children's uncertainty monitoring have documented that older children are typically less confident on incorrect answers than younger children are, but there are relatively similar levels of confidence on correct answers across age-groups in the studies (Hembacher & Ghetti, 2014; Lyons & Ghetti, 2011). Moreover, a difference score between accurate and inaccurate confidence levels has provided an intuitive index of individual difference in uncertainty monitoring. However, these types of metrics do not account for differences in accuracy level in the cognitive task and do not account for other response biases (e.g., a tendency to indiscriminately claim higher levels of confidence) which can both influence measures of uncertainty monitoring (Fleming & Lau, 2014; Rahnev & Fleming, 2019). Newer methods have addressed these limitations (Fleming & Lau, 2014).

One increasingly popular approach in research with adults is to apply a signal detection theory (SDT) framework to accuracy and confidence data (Fleming & Lau, 2014) and calculate parameter estimates, such as d' for cognitive performance and meta-d' for metacognitive performance (Maniscalco & Lau, 2012, 2014). As is the case for simple differences in average confidence, metrics within the SDT framework such as meta-d' estimate the degree to which individuals' confidence ratings distinguish between correct and incorrect responses, while accounting for accuracy levels and response biases on the cognitive task. For example, higher confidence ratings on a correctly identified trial and lower confidence on a false alarm would lead to a bigger meta-d' than indiscriminately high confidence across both types of trials. A popular application of these metrics involves calculating the *m*-ratio (i.e., meta-d'/d'), which can be conceptualized as the level of metacognitive efficiency for a given amount of available cognitive signal for children to monitor (Maniscalco & Lau, 2012). In this sense, an *m*-ratio of 1 represents optimal metacognitive efficiency (and would correspond conceptually high mean difference scores between confidence for correct versus incorrect responses in the traditional methods; Fleming & Lau, 2014) and lesser values can be seen as a percentage of the available signal that children were able to monitor.

A drawback of using this approach, however, is that hundreds of trials are typically necessary to obtain stable estimates at the individual participant level, which can limit their utility in studies of individual differences using special populations like developmental samples with young children. However, recent advances have provided a way to calculate meta-d' and the *m*-ratio indices at the individual level by relying on hierarchical methods within a Bayesian framework (Fleming, 2017) with significantly fewer trials per participant which is especially well suited for studies of young children. We took this approach in the present research to characterize the developmental change in uncertainty monitoring during the kindergarten year.

Correlates and predictors of uncertainty monitoring

The development of uncertainty monitoring does not occur in a vacuum. In the current research, we sought to characterize the relation between uncertainty monitoring and two relevant constructs, namely executive function, and vocabulary development.

Executive function

Children's executive function involves a set of cognitive skills that support changes in children's ability to effectively navigate their world (McClelland & Cameron, 2012; Zelazo et al., 2003). Executive function and its behavioral manifestations are typically characterized as a complex construct consisting of several related subcomponents, such as inhibitory control, attention shifting, and working memory (Garon et al., 2008; Miyake & Friedman, 2012). Young children's executive function undergoes dramatic changes in both capacity and organization during early childhood (Lee et al., 2013) which has implications for broad range of important behavioral and academic outcomes (McClelland, Cameron, Connor, et al., 2007; McClelland, Cameron, Wanless, et al., 2007; McClelland et al., 2014; Schmitt et al., 2017).

Executive function in early childhood is thought to provide at least part of the foundation for self-monitoring and behavioral regulation (Lyons & Zelazo, 2011; Roebers, 2017; Zelazo et al., 2018), but the exact relation between children's executive function and uncertainty monitoring is not well understood. Empirical work connecting the two constructs in young children has been rare (but see Spiess et al., 2016 for a recent example) despite theoretical frameworks highlighting shared developmental time course, neurocognitive structures, and predictive outcomes (Roebers, 2017). Thus, more research is necessary to document the relation between the two constructs empirically, and to elucidate the prerequisite skills and trajectories associated with development of each skill over time.

One possibility is that children's early developing executive function might act as an initial scaffolding for the development of uncertainty monitoring. Specifically, children's increasing ability to inhibit prepotent behaviors, shift their attention inwards, and hold more information in mind would allow them to better reflect upon their uncertainty states and recognize how these feelings are related to subsequent behavior. Some previous studies have found associations between each of these components and older children's metacognitive skills (e.g., Roebers et al., 2012), which highlights the importance of these skills in middle-childhood. However, reviews of these associations suggest that the relation might diminish over development (Roebers & Feurer, 2016). From this perspective, children's earlier executive function skills may predict later uncertainty monitoring, but further development of executive function might not provide additional unique benefits.

Vocabulary

Children's vocabulary development has strong implications for several areas in cognitive and social development (Cole et al., 2010; Kastner et al., 2001). Children's vocabulary in the preschool years is a consistent predictor of the development of sociocognitive skills, such as the acquisition of an explicit theory of mind (Milligan et al., 2007). To our knowledge, the relation between vocabulary development and early uncertainty monitoring has not been directly examined; however, in studies with older children, different aspects of children's broader metacognitive skillsets such as children's metacognitive knowledge have been found to be related to their language skills (Annevirta et al., 2007; Boulware-Gooden et al., 2007; Ebert, 2015; Lecce et al., 2010). Thus, a similar relation may be also hypothesized for younger children, but the possible mechanisms behind this relation are not well understood.

There has been speculation of a direct causal relation between early children's vocabulary skills and aspects of metacognitive skills, such as uncertainty monitoring (Ebert, 2015). Children with better vocabulary skills might have more defined knowledge structures to better recognize their subjective states including feelings of uncertainty. For example, they might be able to better label and report feelings of uncertainty and this ability may support their learning of associations between feelings of uncertainty and their prior or subsequent behaviors. Parental conversation styles that focus on the use of mental verbs have been found to be related to children's vocabulary and knowledge of the mind (Devine & Hughes, 2019) and a similar relation could extend to children's metacognitive skills. These types of relations would be consistent with relations between children's vocabulary skills and other cognitive domains, such as their understanding of emotional states (Cole et al., 2010). Beyond associations with vocabulary levels, growth in children's vocabulary skills over time might also support the development of uncertainty monitoring. The acquisition of increasingly sophisticated vocabulary skills may support children's gains in the ability to utilize fine distinctions among metacognitive states. In this case, children's growth in vocabulary over time should also lead to better uncertainty monitoring abilities.

It is also possible that the relation between vocabulary and uncertainty monitoring may depend on the pragmatics of uncertainty monitoring tasks. These tasks are verbal in nature and thus may facilitate performance in children with high vocabulary. For example, children with better language skills may understand the task instructions better and/or learn more readily the labels associated with each confidence level. Conversely, verbal tasks may hinder performance in children with lower vocabulary. Finally, children's vocabulary may predict uncertainty monitoring because vocabulary may reflect more domain-general processes. Children's vocabulary scores typically correlate highly with measures of general intelligence (Campbell et al., 2001) and, for this reason, are often used as a proxy for this construct.

Overall, there is a good theoretical and empirical basis to expect that executive function and vocabulary support the development of uncertainty monitoring. For the current study, we had the opportunity to include assessments of uncertainty monitoring in an ongoing longitudinal research project in which both executive function and vocabulary were assessed from preschool into kindergarten, allowing us to test for our stated alternative hypotheses.

The current study

The goal of this study was to evaluate the development and predictors of children's uncertainty monitoring in early childhood with a sample of children from families with low income. We administered the same perceptual judgment task utilized in previous studies of early uncertainty monitoring (Coughlin et al., 2015) in the fall and spring of the kindergarten year. Our predictors of uncertainty monitoring included measures of children's executive function and their receptive and expressive vocabulary. Assessment for these predictors was available not only for Fall and Spring of kindergarten, but also for Fall and Spring of the preceding preschool year.

Evaluating the development of uncertainty monitoring

First, we conducted a confirmatory analysis to verify whether young children from families with low income, which are typically underrepresented in empirical research on metacognition, would demonstrate uncertainty monitoring at the transition to kindergarten as found in previous studies (Coughlin et al., 2015; Hembacher & Ghetti, 2014; Lyons & Ghetti, 2011, 2013). We expected children to demonstrate it by reporting higher confidence on correct judgments compared to incorrect judgments. Second, we tested whether these children would demonstrate significant growth in their uncertainty monitoring when measured in the fall and spring of the kindergarten year. To document longitudinal change, we examined complementary indices of uncertainty monitoring, including the traditional average difference of children's confidence ratings for accurate responses minus confidence ratings for inaccurate responses (e.g., Coughlin et al., 2015) as well as more recent approaches derived from SDT framework (Fleming

& Lau, 2014) estimating parameters of metacognitive efficiency as detailed in the Methods section.

Evaluating predictors of uncertainty monitoring

We conducted exploratory analyses examining predictors of uncertainty monitoring. We expected that children's starting point in performance on a measure of executive function would be related to children's uncertainty monitoring indicating the importance of early aspects of children's executive functioning. Additionally, we tested whether children's growth in performance on the executive function task from preschool to kindergarten predicted uncertainty monitoring. We did not expect growth in children's executive function skills to be uniquely related to their uncertainty monitoring over and above their starting point.

We also assessed children's vocabulary. We expected that the starting point of children's vocabulary skills would be related to uncertainty monitoring indicating the importance of children's ability to recognize and describe their feelings of uncertainty. Additionally, we expected that growth in children's vocabulary skills would be related to uncertainty monitoring indicting that the acquisition of new vocabulary might allow for the continued practice and refinement of children's early uncertainty monitoring ability.

METHOD

Participants

The present study consisted of 137 children (52% female), recruited from 42 classrooms nested in 14 Head Start preschools in the Pacific Northwest (average cluster size = 3.25 children per classroom). Children were followed from fall of preschool (M_{age} = 4.70 years, *SD* = 0.29) to spring of kindergarten (M_{age} = 6.10 years, *SD* = 0.29) in 2018 and 2019. At the start of kindergarten, children were distributed across 69 classrooms across 39 sites (average cluster size = 1.62 children per classroom). Children in this study came from a single cohort of a larger longitudinal study examining the relation between measures of executive function and academic outcomes (McClelland et al., 2021). Approximately 60% of participants returned a parent demographic form while their child was enrolled in the study. The average parent education was 12.16 years (*SD* = 2.17), with 46% reporting a high school education or less. Participants could select multiple racial/ethnic identities; 82% reported "White," 16% reported "Latino/Hispanic," and 2% marked another race/ethnicity. Of these participants, 22% marked two or more races; most frequently reporting Latinx/Hispanic, Pacific Islander, and African American identities. At the time of the study, 37% of the households involved a single parent, and 34% were not employed.

Procedure

Children were assessed individually in a classroom or hallway of children's school away from other children during free play periods. All assessments were administered by trained graduate and undergraduate research assistants. At each timepoint, children provided verbal assent prior to each session, and sessions were kept under 20 min to reduce fatigue. When notified by a caregiver or teacher that a child spoke Spanish primarily at home (20% of sample at fall of preschool, 18% of sample at spring of kindergarten), a bilingual experimenter began each session with an English proficiency screener (the preLAS; Duncan & De Avila, 1998). Children who passed the screener were given the rest of the battery in English. Spanish-speaking children who did not pass the screener were assessed by a bilingual experimenter in Spanish. At fall of preschool, 14% of children were assessed in Spanish; but by spring of Kindergarten, only 3% were assessed in Spanish.

Measures

English proficiency screener

The present study utilized two of the most common subtests within the preLAS (Duncan & De Avila, 1998): "Simon Says" (receptive vocabulary) and "Art Show" (expressive vocabulary) to assess children's English language proficiency. Each subtest is comprised of 10 items, in which children receive one point for a correct response and zero points for an incorrect response. Children scoring at or above 15 of 20 points were assessed in English for all other measures. This cut score aligns with current literature (Rainelli et al., 2017) and is based on the publisher's recommendations (Duncan & De Avila, 2000). Reliabilities ranged from $\alpha = .76$ to $\alpha = .90$ across the four time points.

Uncertainty monitoring

To assess children's uncertainty monitoring, we used the same perceptual judgment task with confidence ratings utilized in previous studies with similar aged-children (Coughlin et al., 2015). In the perceptual judgment task, children were shown 20 trials each presenting two degraded line-drawings depicting different objects and were asked to identify the target object and then rate their confidence. Children chose the target object by pointing to or touching the object they thought was the target. After choosing a picture, children then chose their level of certainty using a three-point pictorial confidence scale depicting a "very-sure," "kind-of-sure," and "not-sure" choices. All line-drawings in the task were degraded by an automated computer program which randomly removed a proportion of pixels from each image. Line-drawings had 40% of all pixels of the image removed which is in line with the degradation levels used with this same set of stimuli in previous studies with similar aged-children (Coughlin et al., 2015; Lyons & Ghetti, 2013). Children were trained on how to use the confidence scale using the exact same procedures utilized in previous studies (Coughlin et al., 2015; Hembacher & Ghetti, 2014; Lyons & Ghetti, 2013) which have proven to be effective in children as young as 3 years of age. Unlike previous studies which used computerized software to present stimuli and record children's responses, a pen and paper version of the task was administered so it could be more easily integrated with usual pedagogical procedures and thus be minimally obtrusive within the classroom. Stimuli pairs were printed out on letter sized paper and presented one at a time in a three-ring binder. Children's responses were recorded by the experimenter on a paper score sheet.

Prior to starting the task, children received two practice trials in which they were asked to identify the target object and the experimenter provided certainty-based feedback (e.g., "you were really sure about that one!"). Children were then introduced and trained on how to use the confidence scale. After training, children received four additional practice trials with confidence ratings and the experimenter provided feedback based on children's use of the confidence scale. (i.e., correcting children if they expressed low confidence but chose a high confidence point on the scale). After completing the practice trials, children received one final prompt on how to use the confidence scale before starting the test trials. No feedback was given during the test trials. Children's accuracy was scored as 0 for an incorrect choice and 1 for a correct choice. Children's confidence ratings were scored as 0 for "not-sure" choices, 1 for "kind-of-sure" choices, and 2 for "very-sure" choices. This task was administered twice, once in the Fall and once in the Spring of Kindergarten, approximately 6 months apart.

Executive function

To assess children's executive functioning skills, we used a newly revised version of the Head–Toes–Knees–Shoulders (HTKS-R) task (Gonzales et al., 2021). The HTKS-R assesses aspects of children's inhibitory control, working memory, and flexibility (McClelland et al., 2014) and is appropriate for children ages 4–8 years old. During the standard task, the children are given a pair of behavioral commands and are asked to do the opposite of what they are told (e.g., if told to touch their head, the child must touch their toes). The task consists of three subsections increasing in complexity. In Part 1, children are given a single pair of behavioral commands to remember (e.g., "Head/Toes"). In part two, children are given an additional pair of behavioral commands to remember (e.g., "Head/Toes" and "Knees/Shoulders"), and in Part 3, the pair of behavioral commands are switched (e.g., "Head/Knees", "Shoulders/Toes"). The HTKS-R also includes a new initial section designed to capture additional variability in early aspects of children's executive functioning skills. The new section consists of verbal commands and responses where children are asked to verbally respond to prompts, such as: "When I say toes, you say head." Throughout all sections of the measure, children receive two points for a correct response, one point for a self-corrected response, and zero points for an incorrect response. Scores range from 0 to 118, and the measure demonstrated high internal consistency. This measure was administered four times across the fall and spring of the preschool and kindergarten year with each assessment approximately 6 months apart (Cronbach's α: fall of Pre-K = .95, spring of Pre-K = .95, fall of K = .94, spring of K = .93).

Vocabulary

Children's expressive and receptive vocabulary was assessed with the Picture Vocabulary subtest of the Woodcock–Johnson-III (Muñoz-Sandoval et al., 2005; Woodcock et al., 2001). Previous studies verify that the English and Spanish versions of the WJ-III subtests measure the same constructs, demonstrate similar levels of difficulty, and scores can be combined for statistical analyses (Woodcock et al., 2001; Woodcock & Muñoz-Sandoval, 1993). The Picture Vocabulary subtest was administered four times across the fall and spring of the preschool and kindergarten year with each assessment approximately 6 months apart Previous research examining the standardized assessments have typically found high reliabilities (α < .80) for all subtests in the battery for both the English (Woodcock et al., 2001) and Spanish versions of the assessments (Muñoz-Sandoval et al., 2005).

Analytic approach

We took complementary analytical approaches to provide convergent evidence for the development of uncertainty monitoring. First, we compared children's average confidence on correct and incorrect trials as in previous studies (Coughlin et al., 2015; Lyons & Ghetti, 2011, 2013) to connect our findings with the literature most directly. Then, we conducted an additional analysis fitting a multi-level model in which we examined how the relation between trial accuracy and children's individual confidence ratings changed over time. This type of analysis maximizes statistical power while also allowing for the examination of how children's use of individual decisions and points in the confidence scale change over time. An additional advantage of this approach is confidence ratings for accurate and inaccurate responses are weighted appropriately (e.g., a child who is accurate in 60% of the trials will necessarily contribute fewer confidence ratings for accurate trials than a child who is accurate in 70% of the trials). Finally, we also assessed individual differences in uncertainty monitoring in an SDT framework using the hMeta-d Toolbox (Fleming, 2017) to calculate individual parameter estimates for *d*′, meta-*d*′, and the *m*-ratio. Meta-*d*′, and the *m*-ratio provide an estimation of children's uncertainty monitoring while accounting for differences in accuracy levels on the cognitive task and any response biases in children's confidence judgments. Similar to *d*′, meta-*d*′ is estimated as a *z*-score which captures the amount of separation between confidence ratings on correct and incorrect responses. A *z*-score of zero represents no separation. We used the *m*-ratio (i.e., meta-*d*′/*d*′) to assess the relation between children's uncertainty monitoring and longitudinal measures of children's executive function and vocabulary with models in an structural equation modeling (SEM) framework.

Nested data

Because some children in the study shared a classroom environment, intra-class correlations were examined to explore classroom-level influences on child performance on direct assessments. Between-classroom variation accounted for more than 5% of the variation for each direct assessment during at least one time point, above acceptable parameters according to field standards (Hox et al., 2010). As such, all models utilized clustered-robust standard errors when applicable to adjust for the higher-order nested structure of children nested within different classrooms.

Mixed effects models

We used a mixed-effects (i.e., multilevel) ordinal logistic regression model to examine children's confidence and accuracy data at the trial level. This type of model can account for the ordinal nature of children's confidence judgments and for individual trials (Level-1) being nested within children (Level-2). Children's confidence ratings for each trial were treated as the dependent variable, and the trial accuracy (correct = 1) was included as a predictor in Level-1 of the model. The time point of assessment was dummy coded (Spring of Kindergarten = 1) and included as a Level-2 variable. To assess whether children demonstrated a significant increase in the difference in confidence between correct and incorrect trials across the two timepoints, we included a cross-level interaction term between children's accuracy predicting confidence and time point. A significant cross-level interaction term between accuracy and time point in this model would indicate that the relation between children's accuracy and their confidence ratings was different between the two time points. This interaction can be further probed to test whether children are more or less likely to endorse high or low confidence ratings at each level of accuracy across the two time points.

Modeling change in SDT metrics

Using a hierarchical Bayesian approach with the hMeta-d toolbox (Fleming, 2017), we estimated individual uncertainty monitoring parameters for each of the two time points. Specifically, this method utilizes data from the entire sample across a single time point or multiple time points and extrapolates from this estimation to calculate metrics at the individual level. For analyses that made direct mean comparisons across timepoints, we included data from both timepoints simultaneously to generate parameter estimates for each timepoint individually. This ensure that values from the fall and spring of kindergarten timepoints were extrapolated from the same set of estimates and thus would be directly comparable. However, this method requires listwise deletion for missing data at either timepoint. Thus, meta-*d*′ and the *m*-ratio could only be calculated for the 95 participants with complete data on the uncertainty monitoring task. For analyses that did not make direct mean comparisons, we generated parameter estimates from each timepoint individually so that estimates could be created for the full sample of available data. This resulted in parameter estimates being calculated for 100 participants in the fall of kindergarten, and 106 participants in the spring of kindergarten.

Structural equation models

To assess the longitudinal relations between children's executive functioning, vocabulary skills, and children's uncertainty monitoring, we first assessed children's growth in executive functioning and vocabulary in linear growth models in an SEM framework using M-Plus version 8.4 (Muthén & Muthén, 2010). In a structural equation model, we first fit an unconditional linear growth model estimating a latent intercept and slope parameter from all available time points across the preschool and kindergarten years. For the latent intercept parameter in each model, the factor loading for all time points was set to one. For the latent slope parameter in each model, we set the factor loading for children's scores at the first time point in the fall of preschool to zero and incremented the factor loading for each subsequent time point by one. Additionally, we constrained all latent means of the observed variables to zero and the residual variances to be equal over time. Model fit was assessed against an intercept only model where the factor loadings for the latent slope variable were all fixed to one (i.e., a no growth model). Model fit was also assessed via relative model fit for the linear growth model for each variable. Given the modest sample size for these analyses, we expected indices of relative model fit to only be in the moderate range (Iacobucci, 2010).

After fitting and assessing growth models for each predictor variable, we implemented a full structural equation model predicting children's uncertainty monitoring at the end of kindergarten from their latent slope and intercept parameters of each predictor variable while controlling for their uncertainty monitoring in the Fall of kindergarten. With each model, we included additional demographic covariates such as children's age, gender, ELL (English language learner) status, and parent education levels. With this approach, we could assess how children's starting point and growth in each outcome variable were associated with children's uncertainty monitoring at the end of kindergarten.

RESULTS

Descriptive statistics and missing data

Descriptive statistics for all variables in the study are presented in Table 1, and bivariate correlations between the variables are reported in Table 2. Missingness fluctuated between 8% and 28% for direct assessments, and between 4% and 22% for demographic variables. Missing data within each variable were either due to attrition between time points, children being absent during one of the scheduled visits for data collection within a time point, or children refusing to complete a measure during a visit. Missingness was not consistently related to any demographic variables or direct assessments; as such data were assumed to be missing at random (MAR). When applicable, full information maximum likelihood (FIML) estimators were utilized in analyses described below to account for patterns of missing data under MAR assumption (Enders, 2010).

Evidence of uncertainty monitoring

Evaluating the development of uncertainty monitoring

We first assessed whether children demonstrated evidence of uncertainty monitoring by examining their

TABLE 1 Descriptive statistics

	Fall Pre-K			Spring Pre-K			Fall K			Spring K		
	N	*M*	*SD*	*N*	*M*	*SD*	*N*	*M*	*SD*	*N*	*M*	*SD*
Average accuracy	—	—	—	—	—	—	100	0.65	0.11	106	0.69	0.12
Average incorrect certainty	—	—	—	—	—	—	100	1.35	0.41	106	1.39	0.45
Average correct certainty	—	—	—	—	—	—	100	1.46	0.34	106	1.56	0.34
Average certainty difference	—	—	—	—	—	—	100	0.10	0.32	106	0.17	0.35
d'	—	—	—	—	—	—	95	0.78	0.69	95	1.02	0.60
Meta-*d'*	—	—	—	—	—	—	95	0.33	0.60	95	0.46	0.31
m-ratio	—	—	—	—	—	—	95	0.40	0.23	95	0.46	0.08
HTKS	127	36.57	30.73	99	51.70	36.88	113	67.97	33.89	109	86.40	29.01
WJ-Picture Vocabulary	129	457.18	18.69	97	465.45	14.92	112	467.79	15.72	109	471.72	11.88
Age (months)	132	56.33	3.45	115	61.40	3.51	113	67.11	3.45	108	73.07	3.46
Gender (% male)	—	48%	—	—	47%	—	—	45%	—	—	46%	—
% ELL	—	14%	—	—	13%	—	—	8%	—	—	3%	—

Abbreviation: ELL, English Language Learner; HTKS, Head–Toes–Knees–Shoulders; WJ, Woodcock–Johnson.

TABLE 2 Bivariate correlations

	1	2	3	4	5	6	7	8	9	10	11	12	13
1. *m*-ratio Fall-K	—												
2. *m*-ratio Spring-K	−.07	—											
3. HTKS Fall Pre-K	.03	.18	—										
4. HTKS Spring Pre-K	−.06	.17	.73***	—									
5. HTKS Fall-K	.10	.22*	.70***	.80***	—								
6. HTKS Spring-K	.11	.19*	.61***	.56***	.73***	—							
7. WJ-PV Fall Pre-K	−.01	.08	.43***	.35**	.29**	.24**	—						
8. WJ-PV Spring Pre-K	−.07	.16	.48***	.40***	.26*	.24*	.60***	—					
9. WJ-PV Fall-K	.02	.24*	.59***	.37***	.48***	.51***	.53***	.61***	—				
10. WJ-PV Spring-K	.02	.25*	.56***	.39***	.51***	.47***	.40***	.46***	.61***	—			
11. Age	.07	.13	.32***	.13	.30**	.27**	.30**	.31**	.36***	.25**	—		
12. Gender	.17	.09	−.02	−.10	−.06	−.04	−.08	−.13	−.13	−.11	.06	—	
13. ELL status	−.03	−.14	−.28**	−.12	−.29**	−.30**	−.20***	−.08***	−.43***	−.47***	−.12	.10	—
14. Parent education	.04	−.08	.26*	.31*	.22	.18	.23	.28*	.14	.31*	.02	.01	−.41***

Abbreviations: ELL, English Language Learner; HTKS, Head–Toes–Knees–Shoulders; WJ-PV, Woodcock–Johnson Picture Vocabulary.
*$p < .05$.
**$p < .01$.
***$p < .001$.

average confidence ratings on accurate trials and their average confidence on inaccurate trials during the fall and spring of the kindergarten year. As shown in Figure 1a, children indicated higher confidence on correct trials (M = 1.46, SD = 0.34) than on incorrect trials (M = 1.35, SD = 0.41) in the Fall of kindergarten, $t(99) = 3.25$, $p = .002$, Cohen's $d = .33$, and as shown in Figure 1b, children also indicated higher confidence on correct (M = 1.56, SD = 0.34) versus incorrect (M = 1.39, SD = 0.45) trials in the spring of kindergarten as well, $t(105) = 5.09$, $p < .001$, Cohen's $d = .50$. These differences are similar in magnitude to other studies examining confidence rating differences in similar-aged children using the same paradigm (Coughlin et al., 2015; Lyons & Ghetti, 2013). When we examined the raw differences in children's average confidence ratings, no significant difference emerged between the Fall (M = 0.11, SD = 0.32) and Spring (M = 0.18, SD = 0.36) of the kindergarten year, $t(94) = 1.37$, $p = .175$. Children's overall accuracy did improve between the Fall (M = 0.65, SD = 0.11) and Spring (M = 0.70, SD = 0.11) of the kindergarten year, $t(94) = 3.17$, $p = .002$, Cohen's $d = .39$.

Mixed effects models

When examining children's confidence ratings at the trial level, there was a significant main effect of trial accuracy on children's confidence judgments ($\beta = .30$, $p < .001$, OR = 1.35 [95% CI: 1.15–1.59]), which is conceptually analogous to children's mean average confidence differences between accurate and inaccurate responses reported above. There was no main effect of time point ($\beta = .20$, $p = .19$, OR = 1.22 [95% CI: 0.91–1.63]) indicating there were no differences in children's overall confidence judgments over time; however, there was a significant cross-level interaction term between the effect of accuracy and time point on children's confidence ratings ($\beta = .27$, $p = .04$, OR = 1.31 [95% CI: 1.00–1.71]). This significant interaction indicates that the relation between children's confidence ratings and trial accuracy changed between the Fall of kindergarten and the Spring of kindergarten. Probing this interaction further revealed that increases in uncertainty monitoring over time came from children being more likely to choose "very sure" response on the confidence scale on correct trials ($z = 2.50$, $p = .01$) in the spring of kindergarten (probability = .69, 95% CI [0.64–0.73]) versus the fall of kindergarten (probability = .62, 95% CI [0.58–0.66]). Children were also less likely to use the middle "kind of sure" response on correct trials ($z = -3.50$, $p < .001$) across the Fall (probability = .22, 95% CI [0.19–0.25]) and Spring (probability = .18, 95% CI [0.15–0.20]), respectively. Similarly, for incorrect trials, children were also less likely to use the "kind of sure" response ($z = -2.79$, $p = .005$) across the fall (probability = .25, 95% CI [0.21–0.28]) and spring (probability = .21, 95% CI [0.18–0.24]), respectively. There were no significant changes across time in children's use of the "very sure" response on incorrect trials and children's use of the "not so sure" response for either correct or incorrect trials.

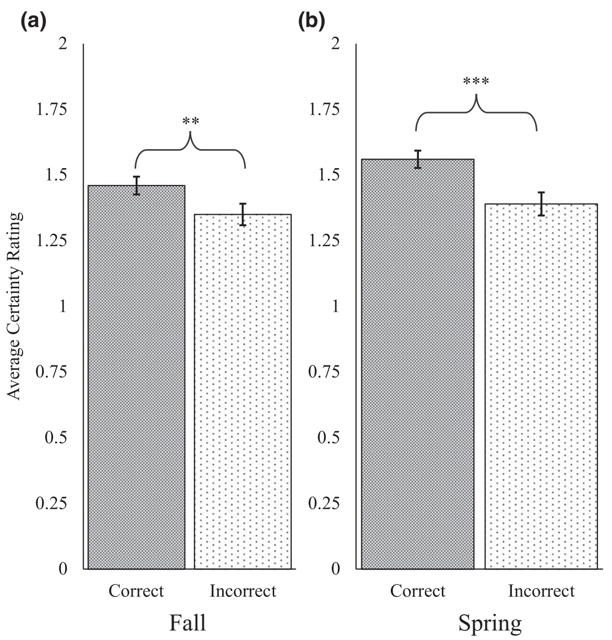

FIGURE 1 Children's average certainty (±1 SE) on correct and incorrect trials in the perceptual judgment task in the (a) Fall of kindergarten and (b) Spring of kindergarten. **$p < .01$, *** $p < .001$

SDT metrics

When examining changes in the SDT metrics across the Fall and Spring of the kindergarten year, there were improvements in children's sensitivity to identify the correct object in the task (d') in the Fall (M = 0.78, SD = 0.69) and Spring (M = 1.02, SD = 0.60), $t(94) = 2.76$, $p = .007$, Cohen's $d = .33$, as well as in their metacognitive sensitivity (meta-d') in the Fall (M = 0.33, SD = 0.60) and Spring (M = 0.46, SD = 0.31), $t(94) = 2.63$, $p = .010$, Cohen's $d = .30$. In addition, there were improvements in children's metacognitive efficiency as indexed by the m-ratio (meta d'/d'). Estimates of children's metacognitive efficiency increased from the fall of kindergarten (M = 0.40, SD = 0.23) compared to the spring of kindergarten (M = 0.46, SD = 0.08), $t(94) = 2.32$, $p = .023$, Cohen's $d = .24$. These results indicate that measuring children's metacognitive skills via a SDT approach within a Bayesian framework can provide additional utility and sensitivity for group-level comparisons compared to simpler metrics.

Individual differences in uncertainty monitoring

When we examined individual differences in uncertainty monitoring, no significant correlation was observed between fall and spring of kindergarten with either difference scores between average confidence on correct and incorrect judgments, $r(93) = .02$, $p = .79$, or with estimates of the m-ratio, $r(93) = -.09$, $p = .36$. Thus, stable individual differences in uncertainty monitoring observed with children in this age group may still be emerging.

Longitudinal relations with uncertainty monitoring

HTKS-R

Changes in children's performance on the HTKS-R were evaluated using a linear growth model fitting separate latent parameters for the intercept and slope. The solution for the linear growth model demonstrated only moderate levels of relative model fit as expected given the available sample size, $\chi^2(8) = 40.26$, $p < .001$, comparative fit index (CFI) = .89, root mean square error of approximation (RMSEA) = .17, however, there was a significant increase in absolute model from the no-growth model to the unconditional linear growth model $\chi^2(3) = 278.83$, $p < .001$. Next, we assessed the relation between the latent parameters for the HTKS and children's uncertainty monitoring in the Spring of Kindergarten while controlling for levels of children's uncertainty monitoring in the Fall of Kindergarten. As shown in Figure 2a, children's starting point on the HTKS-R (i.e., the intercept) predicted children's uncertainty monitoring, $\beta = .33$, $p = .006$, but children's growth on the HTKS-R across the preschool and kindergarten years (i.e., the slope) did not predict uncertainty monitoring, $\beta = .21$, $p = .40$. No other covariates were significant independent predictors of children's uncertainty monitoring. This indicates that children's starting level of executive function at the beginning of preschool uniquely predicted uncertainty monitoring at the end of kindergarten, accounting for initial uncertainty monitoring, but growth in children's executive function over the 2 years did not.

Picture vocabulary

Changes in children's performance on the WJ-Picture Vocabulary task were evaluated using the same type of linear growth model fitting as with the HTKS-R task. The solution for the linear growth model for the WJ-Picture Vocabulary task demonstrated similar moderate levels of relative model fit, $\chi^2(8) = 17.61$, $p = .024$, CFI = .90, RMSEA = .09. Similar to the growth model for the HTKS-R, there was a significant increase in

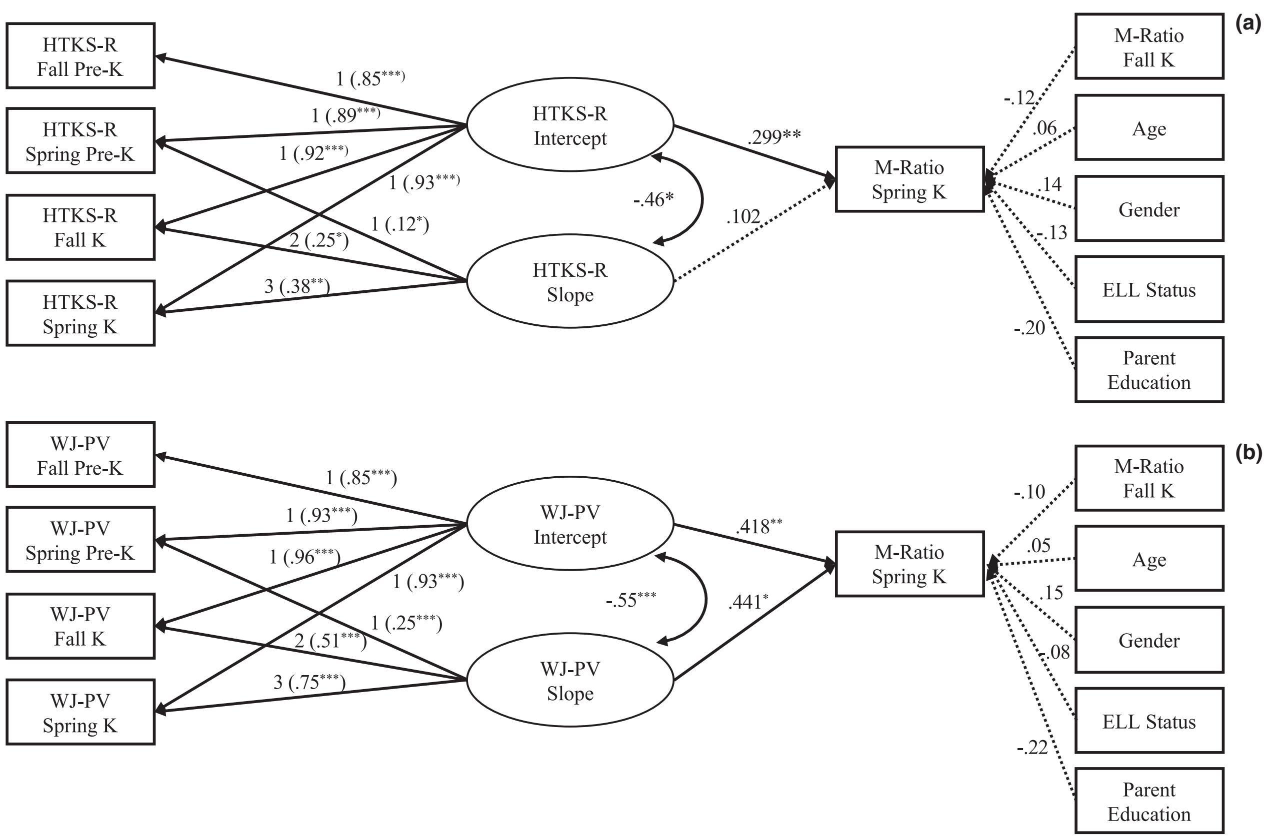

FIGURE 2 Structural equation models depicting the fixed intercept and slope parameters (with standardized factor loadings in parentheses) for assessments at each time point and standardized beta coefficients and correlations between the (a) HTKS-R or (b) WJ-PV and uncertainty monitoring. HTKS-R, Head–Toes–Knees–Shoulders revised; WJ-PV, Woodcock–Johnson Picture Vocabulary. $^*p < .05$, $^{**}p < .01$, $^{***}p < .001$

model fit from the no-growth model to the unconditional linear growth model $\chi^2(3) = 155.38$, $p < .001$. We next assessed how the latent parameters for the WJ-Picture Vocabulary predicted children's uncertainty monitoring. As shown in Figure 2b, children's starting performance level on the WJ-Picture vocabulary task (i.e., the intercept) was related to children's uncertainty monitoring, $\beta = .98$, $p = .02$, and children's growth on the WJ-Picture vocabulary task (i.e., the slope) was also related to children's uncertainty monitoring, $\beta = .93$, $p = .03$. No other covariates were significant independent predictors of children's uncertainty monitoring. This indicates that both children's starting point in their vocabulary as they entered preschool and their growth in vocabulary across the preschool and kindergarten year were both unique and independent predictors of children's metacognitive skills at the end of kindergarten.

DISCUSSION

The goal of the current study was to examine the development of uncertainty monitoring at the transition to kindergarten with a sample of children from economically disadvantaged families. In addition, we investigated predictors of children's uncertainty monitoring incorporating available longitudinal data on measures of children's executive functioning and vocabulary skills from the preschool and kindergarten years.

Uncertainty monitoring in early childhood

Our results provide evidence for the presence of uncertainty monitoring in early childhood which is consistent with the findings from previous studies (Coughlin et al., 2015; Hembacher & Ghetti, 2014; Lyons & Ghetti, 2011, 2013). In both the Fall and Spring of the kindergarten year, children demonstrated significantly higher confidence in accurate trials compared to inaccurate trials in a perceptual judgment task. We also documented significant growth in children's uncertainty monitoring from the Fall of kindergarten to the Spring of kindergarten. These results were revealed using more advanced methods to estimate children's uncertainty monitoring compared to traditional indicators. Modeling children's confidence ratings and trial accuracy at the trial level demonstrated that improvements in children's overall uncertainty monitoring largely came from children being more likely to endorse high confidence on correct trials in Spring of kindergarten compared to the Fall. Additionally, metacognitive metrics derived from a SDT framework, such as meta-d' and the m-ratio (Fleming & Lau, 2014) revealed significant improvements across the Fall and Spring of kindergarten demonstrating the additional utility and sensitivity compared to traditional metrics.

Together, these results provide further evidence for young children's uncertainty monitoring being a robust phenomenon, which is an important step in beginning to demonstrate the generalizability of such findings (Henrich et al., 2010). Although direct comparisons with other types of study populations were not made, the current findings demonstrate children from families with low income can engage in uncertainty monitoring in early childhood. Even though such samples are more likely to face certain hardships that can impact the pace and patterns of development (Blair & Raver, 2012; Sektnan et al., 2010), the expected development of uncertainty monitoring extends beyond children from families with higher income and education levels. However, children in the current study were enrolled in both a Head Start preschool program which emphasizes early academic skills in its curricula and had an initial exposure to more formal education settings during the kindergarten year. Whether and how these types of experiences impact the development of early metacognitive skills in children from both advantaged and disadvantaged backgrounds remains an important area of investigation for future research.

In contrast to these results, we did not find a significant correlation between measurements of uncertainty monitoring across the fall and spring of the kindergarten year. One possible explanation for these findings is that uncertainty monitoring is still emerging for many children in this period of development. To the extent that some re-organization occurs to support the emergence of uncertainty monitoring during the kindergarten year, the relation between measurements over time would be expected to be weak. Indeed, reliable relations across assessment times have been consistently found in children whose uncertainty monitoring skills are more established (e.g., Fandakova et al., 2017; Roebers & Spiess, 2017).

Longitudinal predictors of uncertainty monitoring

In addition to documenting the development of uncertainty monitoring during early childhood, the current study also provided new evidence about longitudinal predictors of this ability. Specifically, we focused on children's executive function and vocabulary skills. The relation between children's executive function and uncertainty monitoring has been posited in previous theoretical discussions (Roebers, 2017). Evidence from the current study supports the notion that aspects of children's executive functioning in early childhood are important for the development of uncertainty monitoring. We predicted that children's ability to control their behaviors, shift their attentional focus between different internal and external demands, and hold more information in mind would allow children to better reflect upon their subjective feelings of uncertainty and to recognize how these feelings are related to previous or subsequent behavior.

In the current study, estimates of children's starting point in executive function in preschool independently predicted children's uncertainty monitoring, but growth in executive function during the preschool and kindergarten years was not found to independently predict children's uncertainty monitoring over and above their starting point. As hypothesized, earlier developing aspects of executive functions might be sufficient for the expression of young children's uncertainty monitoring, but further development might not be a necessary condition for the development of uncertainty monitoring to occur in early childhood. From this viewpoint, early aspects of children's executive function might act as a scaffolding for the development of uncertainty monitoring in early childhood, and the self-monitoring aspects of executive functioning might develop independently of other related skills.

However, it is possible that different or additional relations between executive function and children's uncertainty monitoring might emerge when utilizing different measurement models of children's executive functioning. During early childhood, executive function develops from a more unitary construct to a more complex phenomenon with more distinct subcomponents (Lee et al., 2013). More direct measures of growth in these later developing aspects of executive function (i.e., attention-shifting, working-memory) or models that more explicitly specify the complexities in executive function development might reveal reciprocal relations over time between aspects of executive function and uncertainty monitoring.

The current study utilized the HTKS-R to measure children's executive functioning skills, and the HTKS-R is a complex measure of executive function that captures multiple subcomponents of children's executive abilities (McClelland et al., 2014; Schmitt et al., 2017). The HTKS-R is also often one of the strongest predictors of many different child outcomes (McClelland et al., 2014, 2021; Wanless et al., 2011), but a latent variable approach that utilizes many measures of children's executive function to explicitly model the different subcomponents as they develop might better elucidate the relation between children's executive function and uncertainty monitoring over time. Thus, future research can better elucidate how children's early inhibitory control skills and later developing working memory relate to the development of children's uncertainty, as well as explore possible reciprocal relations between the development of these two constructs over time.

In addition, examining executive function, we also investigated the relation between children's uncertainty monitoring and vocabulary development. Both children's initial vocabulary level at the start of the preschool year and children's vocabulary growth over the course of the preschool and kindergarten year independently predicted children's uncertainty monitoring at the end of the kindergarten year. This pair of findings suggests several possible relations between children's uncertainty monitoring and vocabulary development.

As in other areas of research in cognitive development, there might be some baseline knowledge structures that are necessary for children's uncertainty monitoring to be expressed. One possibility is that children with better vocabulary skills might be able to better recognize and describe their feelings of uncertainty. This relation might utilize a similar mechanism as the common association found between children's vocabulary and their emotion understanding (Cole et al., 2010). Future research could investigate this relation and help differentiate between alternative explanations by examining specific knowledge structures such as children's mental state vocabulary and children's uncertainty monitoring rather than examining children's general vocabulary skills alone. In related areas of cognitive research, such as children's theory of mind, children's executive function and vocabulary are often unique predictors of children's performance (Carlson et al., 2004). Moreover, vocabulary is associated with both children's and parents' mental state speech and children's theory of mind development (Devine & Hughes, 2018). Specific conversational styles between parents and children that place emphasis on mental verbs describing uncertainty might be particularly beneficial for the development of children's uncertainty monitoring. Future research should assess and differentiate these different aspects of vocabulary to clarify what is most critical for the development of uncertainty monitoring in young children.

The unique relation between growth in children's vocabulary skills and their uncertainty monitoring also allows for the potential of a reciprocal relation wherein children can practice their uncertainty monitoring skills when expanding their vocabulary. In older children, the relation between metacognitive skills and vocabulary have been reported (Annevirta et al., 2007; Boulware-Gooden et al., 2007; Ebert, 2015; Lecce et al., 2010). With young children, uncertainty monitoring could help children recognize when they do or do not know the correct label for an object or meaning of a word they hear in their environment. Children's understanding of their own knowledge and ignorance helps guide them on when to seek information from an adult or more knowledgeable peer (Ronfard et al., 2017). The significant relation between growth in vocabulary and their uncertainty monitoring in the current study supports this notion and indicates that relations between children's metacognitive skills and self-regulated learning can be extended even into the early childhood period of development. As children gain a more advanced vocabulary, they might be able to recognize and make distinctions between more granular levels of uncertainty which could further strengthen their uncertainty monitoring abilities.

Before concluding, we note a few limitations of the current study and potential future directions. First, whereas we provide evidence of uncertainty monitoring in a sample of young children from families with low-income backgrounds, we did not examine the relation between income and uncertainty monitoring directly. It remains possible that income-related differences exist despite evidence of uncertainty monitoring in early childhood across income levels. The examination of associations between income and uncertainty monitoring, including the factors that might explain these associations, remains an important area of future research. Additionally, it remains important to assess potential reciprocal relations between growth in uncertainty monitoring and growth in the other variables. Reciprocal relations between changes in these skills over time are plausible and should be further explored, especially with older children where these skills are more established. Moreover, our relatively limited sample size prevented us from testing an overarching model in which the initial level and growth of both executive function and vocabulary were all included. This limited sample size was also a likely reason for the moderate levels of relative model fit in some of our analyses, which calls for future replication of the current findings. Finally, the uncertainty monitoring task was not delivered to children using the typical laboratory equipment (e.g., touch screen monitor to collect reaction-times or eye-tracker to collect fixations and saccades between stimuli); therefore, we were unable to collect response times or assess evidence accumulation processes from eye movements (Leckey et al., 2020), which have proved informative to characterize the emergence of uncertainty monitoring.

CONCLUSIONS

Overall, findings in the current study provide further evidence that uncertainty monitoring is a robust phenomenon during early childhood and that the development of uncertainty monitoring can be observable in children from both advantaged and disadvantaged backgrounds when utilizing sensitive analytic techniques. Furthermore, the current study provides new evidence for the longitudinal relation between children's uncertainty monitoring and other important constructs such as the development of young children's executive function and vocabulary skills, underscoring relations a broad set of behaviors important to formal educational contexts and self-regulated learning. These findings can help inform the broader literature on aims that seek to both promote and improve positive child outcomes.

ACKNOWLEDGMENTS

Christopher R Gonzales was supported by the Eunice Kennedy Shriver National Institute of Child Health and Human Development of the National Institute of Health through Grant 1F32HD097896-01 awarded to the University of California—Davis (Sponsor: Simona Ghetti). The opinions expressed are those of the authors and do not represent views of the National Institute of Health. The research reported here was supported by the U.S. Department of Education Institute for Education Sciences grants #R305A150192 (PI: McClelland) to Oregon State University. The content is the responsibility of the authors and does not necessarily represent the official views of the Institute of Education Sciences, or the U.S. Department of Education.

ORCID

Christopher R. Gonzales https://orcid.org/0000-0003-4312-5914
Megan M. McClelland https://orcid.org/0000-0002-5604-0541

REFERENCES

Annevirta, T., Laakkonen, E., Kinnunen, R., & Vauras, M. (2007). Developmental dynamics of metacognitive knowledge and text comprehension skill in the first primary school years. *Metacognition and Learning*, *2*(1), 21–39. https://doi.org/10.1007/s11409-007-9005-x

Blair, C., & Raver, C. C. (2012). Child development in the context of adversity: Experiential canalization of brain and behavior. *The American Psychologist*, *67*(4), 309–318. https://doi.org/10.1037/a0027493

Boulware-Gooden, R., Carreker, S., Thornhill, A., & Joshi, R. M. (2007). Instruction of metacognitive strategies enhances reading comprehension and vocabulary achievement of third-grade students. *The Reading Teacher*, *61*(1), 70–77. https://doi.org/10.1598/RT.61.1.7

Campbell, J. M., Bell, S. K., & Keith, L. K. (2001). Concurrent validity of the Peabody Picture Vocabulary Test-third edition as an intelligence and achievement screener for low SES African American children. *Assessment*, *8*(1), 85–94. https://doi.org/10.1177/107319110100800108

Carlson, S. M., Moses, L. J., & Claxton, L. J. (2004). Individual differences in executive functioning and theory of mind: An investigation of inhibitory control and planning ability. *Journal of Experimental Child Psychology*, *87*(4), 299–319. https://doi.org/10.1016/j.jecp.2004.01.002

Cole, P. M., Armstrong, L. M., & Pemberton, C. K. (2010). The role of language in the development of emotion regulation. In S. D. Calkins & M. A. Bell (Eds.), *Human brain development. Child development at the intersection of emotion and cognition* (pp. 59–77). American Psychological Association. https://doi.org/10.1037/12059-004

Coughlin, C., Hembacher, E., Lyons, K. E., & Ghetti, S. (2015). Introspection on uncertainty and judicious help-seeking during the preschool years. *Developmental Science*, *18*(6), 957–971. https://doi.org/10.1111/desc.12271

Devine, R. T., & Hughes, C. (2018). Family correlates of false belief understanding in early childhood: A meta-analysis. *Child Development*, *89*(3), 971–987. https://doi.org/10.1111/cdev.12682

Devine, R. T., & Hughes, C. (2019). Let's talk: Parents' mental talk (not mind-mindedness or mindreading capacity) predicts children's false belief understanding. *Child Development*, *90*(4), 1236–1253. https://doi.org/10.1111/cdev.12990

Duncan, S. E., & De Avila, E. A. (1998). *Pre-Language Assessment Scale 2000*. CTB/McGraw-Hill.

Duncan, S. E., & De Avila, E. A. (2000). *PreLAS 2000 examiner's manual, English forms C & D*. CTB/McGraw-Hill.

Dunlosky, J., & Metcalfe, J. (2008). *Metacognition*. SAGE Publications. https://play.google.com/store/books/details?id=AyNzAwAAQBAJ

Dunlosky, J., & Rawson, K. A. (2012). Overconfidence produces underachievement: Inaccurate self evaluations undermine students' learning and retention. *Learning and Instruction*, *22*(4), 271–280. https://doi.org/10.1016/j.learninstruc.2011.08.003

Ebert, S. (2015). Longitudinal relations between theory of mind and metacognition and the impact of language. *Journal of Cognition and Development*, *16*(4), 559–586. https://doi.org/10.1080/15248372.2014.926272

Enders, C. (2010). *Applied missing data analysis*. Guilford Press.

Fandakova, Y., Selmeczy, D., Leckey, S., Grimm, K. J., Wendelken, C., Bunge, S. A., & Ghetti, S. (2017). Changes in ventromedial prefrontal and insular cortex support the development of metamemory from childhood into adolescence. *Proceedings of the National Academy of Sciences of the United States of America*, *114*(29), 7582–7587. https://doi.org/10.1073/pnas.1703079114

Flavell, J. H. (1979). Metacognition and cognitive monitoring: A new area of cognitive–developmental inquiry. *The American Psychologist*, *34*(10), 906–911. https://doi.org/10.1037/0003-066x.34.10.906

Fleming, S. M. (2017). HMeta-d: Hierarchical Bayesian estimation of metacognitive efficiency from confidence ratings. *Neuroscience of Consciousness*, *2017*(1), 1–14. https://doi.org/10.1093/nc/nix007

Fleming, S. M., & Lau, H. C. (2014). How to measure metacognition. *Frontiers in Human Neuroscience*, *8*(July), 1–9. https://doi.org/10.3389/fnhum.2014.00443

Garon, N., Bryson, S. E., & Smith, I. M. (2008). Executive function in preschoolers: A review using an integrative framework. *Psychological Bulletin*, *134*(1), 31–60. https://doi.org/10.1037/0033-2909.134.1.31

Ghetti, S., Hembacher, E., & Coughlin, C. (2013). Feeling uncertain and acting on it during the preschool years: A metacognitive approach. *Child Development Perspectives*, *7*(3), 160–165. https://doi.org/10.1111/cdep.12035

Gonzales, C. R., Bowles, R., Geldhof, G. J., Cameron, C., Tracy, A., & McClelland, M. M. (2021). The Head-Toes-Knees-Shoulders revised (HTKS-R): Development and psychometric properties of a revision to reduce floor effects. *Early Childhood Quarterly*. https://doi.org/10.1016/j.ecresq.2021.03.008

Hembacher, E., & Ghetti, S. (2014). Don't look at my answer: Subjective uncertainty underlies preschoolers' exclusion of their least accurate memories. *Psychological Science*, *25*(9), 1768–1776. https://doi.org/10.1177/0956797614542273

Henrich, J., Heine, S. J., & Norenzayan, A. (2010). The weirdest people in the world? *The Behavioral and Brain Sciences*, *33*(2–3), 61–83. https://doi.org/10.1017/S0140525X0999152X

Hox, J. J., Moerbeek, M., & van de Schoot, R. (2010). *Multilevel analysis*. Routledge.

Iacobucci, D. (2010). Structural equations modeling: Fit Indices, sample size, and advanced topics. *Journal of Consumer Psychology*, *20*(1), 90–98. https://doi.org/10.1016/j.jcps.2009.09.003

Kastner, J. W., May, W., & Hildman, L. (2001). Relationship between language skills and academic achievement in first grade. *Perceptual and Motor Skills*, *92*(2), 381–390. https://doi.org/10.2466/pms.2001.92.2.381

Koriat, A., & Goldsmith, M. (1996). Monitoring and control processes in the strategic regulation of memory accuracy. *Psychological Review*, *103*(3), 490–517. https://doi.org/10.1037/0033-295X.103.3.490

Lecce, S., Zocchi, S., Pagnin, A., Palladino, P., & Taumoepeau, M. (2010). Reading minds: The relation between children's mental state knowledge and their metaknowledge about reading. *Child Development*, *81*(6), 1876–1893. https://doi.org/10.1111/j.1467-8624.2010.01516.x

Leckey, S., Selmeczy, D., Kazemi, A., Johnson, E. G., Hembacher, E., & Ghetti, S. (2020). Response latencies and eye gaze provide insight on how toddlers gather evidence under uncertainty. *Nature Human Behaviour*, *4*(9), 928–936. https://doi.org/10.1038/s41562-020-0913-y

Lee, K., Bull, R., & Ho, R. M. H. (2013). Developmental changes in executive functioning. *Child Development*, *84*(6), 1933–1953. https://doi.org/10.1111/cdev.12096

Lipko, A. R., Dunlosky, J., & Merriman, W. E. (2009). Persistent overconfidence despite practice: The role of task experience in preschoolers' recall predictions. *Journal of Experimental Child Psychology*, *103*(2), 152–166. https://doi.org/10.1016/j.jecp.2008.10.002

Lockl, K., & Schneider, W. (2007). Knowledge about the mind: Links between theory of mind and later metamemory. *Child Development*, *78*(1), 148–167. https://doi.org/10.1111/j.1467-8624.2007.00990.x

Lyons, K. E., & Ghetti, S. (2010). Metacognitive development in early childhood: New questions about old assumptions. In *Trends and prospects in metacognition research* (pp. 259–278). Springer. https://doi.org/10.1007/978-1-4419-6546-2_12

Lyons, K. E., & Ghetti, S. (2011). The development of uncertainty monitoring in early childhood. *Child Development*, *82*(6), 1778–1787. https://doi.org/10.1111/j.1467-8624.2011.01649.x

Lyons, K. E., & Ghetti, S. (2013). I don't want to pick! Introspection on uncertainty supports early strategic behavior. *Child Development*, *84*(2), 726–736. https://doi.org/10.1111/cdev.12004

Lyons, K. E., & Zelazo, P. D. (2011). Monitoring, metacognition, and executive function. In *Elucidating the role of self-reflection in the development of self-regulation* (1st ed., Vol. *40*, pp. 379–412). Elsevier Inc. https://doi.org/10.1016/B978-0-12-386491-8.00010-4

Maniscalco, B., & Lau, H. (2012). A signal detection theoretic approach for estimating metacognitive sensitivity from confidence ratings. *Consciousness and Cognition*, *21*(1), 422–430. https://doi.org/10.1016/j.concog.2011.09.021

Maniscalco, B., & Lau, H. (2014). Signal detection theory analysis of type 1 and type 2 data: Meta-d′, response-specific meta-d′, and the unequal variance SDT model. *The Cognitive Neuroscience of Metacognition*. https://doi.org/10.1007/978-3-642-45190-4_3

McClelland, M. M., & Cameron, C. E. (2012). Self-regulation early childhood: Improving conceptual clarity and developing ecologically valid measures. *Child Development Perspectives*, *6*(2), 136–142. https://doi.org/10.1111/j.1750-8606.2011.00191.x

McClelland, M. M., Cameron, C. E., Connor, C. M. D., Farris, C. L., Jewkes, A. M., & Morrison, F. J. (2007). Links between behavioral regulation and preschoolers' literacy, vocabulary, and math skills. *Developmental Psychology*, *43*(4), 947–959. https://doi.org/10.1037/0012-1649.43.4.947

McClelland, M. M., Cameron, C. E., Duncan, R., Bowles, R. P., Acock, A. C., Miao, A., & Pratt, M. E. (2014). Predictors of early growth in academic achievement: The head-toes-knees-shoulders task. *Frontiers in Psychology*, *5*(JUN), 1–14. https://doi.org/10.3389/fpsyg.2014.00599

McClelland, M. M., Cameron, C. E., Wanless, S. B., Murray, A., Saracho, O., & Spodek, B. (2007). Executive function, behavioral self-regulation, and social-emotional competence. *Contemporary Perspectives on Social Learning in Early Childhood Education*, *1*, 113–137.

McClelland, M. M., Gonzales, C. R., Cameron, C. E., Geldhof, G. J., Bowles, R. P., Nancarrow, A. F., Tracy, A. (2021). The Head-Toes-Knees-Shoulders revised: Links to academic outcomes and measures of EF in young children. *Frontiers in Psychology*, *12*. https://doi.org/10.3389/fpsyg.2021.721846

Milligan, K., Astington, J. W., & Dack, L. A. (2007). Language and theory of mind: Meta-analysis of the relation between language ability and false-belief understanding. *Child Development*, *78*(2), 622–646. https://doi.org/10.1111/j.1467-8624.2007.01018.x

Miyake, A., & Friedman, N. P. (2012). The nature and organization of individual differences in executive functions: Four general conclusions. *Current Directions in Psychological Science*, *21*(1), 8–14. https://doi.org/10.1177/0963721411429458

Muñoz-Sandoval, A. F., Woodcock, R. W., McGrew, K. S., & Mather, N. (2005). *Bateria III Woodcock-Muñoz: Pruebas de aprovechamiento [Tests of achievement]*. Riverside.
Muthén, L. K., & Muthén, B. O. (2010). *Mplus user's guide* (6th ed.). Muthén & Muthén.
Nielsen, M., Haun, D., Kärtner, J., & Legare, C. H. (2017). The persistent sampling bias in developmental psychology: A call to action. *Journal of Experimental Child Psychology*, *162*, 31–38. https://doi.org/10.1016/j.jecp.2017.04.017
Rahnev, D., & Fleming, S. M. (2019). How experimental procedures influence estimates of metacognitive ability. *Neuroscience of Consciousness*, *2019*(1), 1–42. https://doi.org/10.1093/nc/niz009
Rainelli, S., Bulotsky-Shearer, R. J., Fernandez, V. A., Greenfield, D. B., & Lopez, M. (2017). Validity of the first two subtests of the preschool language assessment scale as a language screener for Spanish-speaking preschool children. *Early Childhood Research Quarterly*, *38*, 10–22. https://doi.org/10.1016/j.ecresq.2016.08.001
Roberts, S. O., Bareket-Shavit, C., Dollins, F. A., Goldie, P. D., & Mortenson, E. (2020). Racial inequality in psychological research: Trends of the past and recommendations for the future. *Perspectives on Psychological Science*, *15*(6), 1295–1309. https://doi.org/10.1177/1745691620927709
Roebers, C. M. (2017). Executive function and metacognition: Towards a unifying framework of cognitive self-regulation. *Developmental Review*, *45*, 31–51. https://doi.org/10.1016/j.dr.2017.04.001
Roebers, C. M., Cimeli, P., Röthlisberger, M., & Neuenschwander, R. (2012). Executive functioning, metacognition, and self-perceived competence in elementary school children: An explorative study on their interrelations and their role for school achievement. *Metacognition and Learning*, *7*(3), 151–173. https://doi.org/10.1007/s11409-012-9089-9
Roebers, C. M., & Feurer, E. (2016). Linking executive functions and procedural metacognition. *Child Development Perspectives*, *10*(1), 39–44. https://doi.org/10.1111/cdep.12159
Roebers, C. M., & Spiess, M. (2017). The development of metacognitive monitoring and control in second graders: A short-term longitudinal study. *Journal of Cognition and Development*, *18*(1), 110–128. https://doi.org/10.1080/15248372.2016.1157079
Roebers, C. M., von der Linden, N., Schneider, W., & Howie, P. (2007). Children's metamemorial judgments in an event recall task. *Journal of Experimental Child Psychology*, *97*(2), 117–137. https://doi.org/10.1016/j.jecp.2006.12.006
Ronfard, S., Bartz, D. T., Cheng, L., Chen, X., & Harris, P. L. (2017). *Children's developing ideas about knowledge and its acquisition* (1st ed.). Elsevier Inc. https://doi.org/10.1016/bs.acdb.2017.10.005
Rowley, S. J., & Camacho, T. C. (2015). Increasing diversity in cognitive developmental research: Issues and solutions. *Journal of Cognition and Development*, *16*(5), 683–692. https://doi.org/10.1080/15248372.2014.976224
Schmitt, S. A., Geldhof, G. J., Purpura, D. J., Duncan, R., & McClelland, M. M. (2017). Examining the relations between executive function, math, and literacy during the transition to kindergarten: A multi-analytic approach. *Journal of Educational Psychology*, *109*(8), 1120–1140. https://doi.org/10.1037/edu0000193
Schmitt, S. A., Purpura, D. J., & Elicker, J. G. (2019). Predictive links among vocabulary, mathematical language, and executive functioning in preschoolers. *Journal of Experimental Child Psychology*, *180*, 55–68. https://doi.org/10.1016/j.jecp.2018.12.005
Sektnan, M., McClelland, M. M., Acock, A., & Morrison, F. J. (2010). Relations between early family risk, children's behavioral regulation, and academic achievement. *Early Childhood Research Quarterly*, *25*(4), 464–479. https://doi.org/10.1016/j.ecresq.2010.02.005
Selmeczy, D., Kazemi, A., & Ghetti, S. (2021). Developmental differences in subjective recollection and its role in decision making. *Child Development*. https://doi.org/10.1111/cdev.13611
Spiess, M. A., Meier, B., & Roebers, C. M. (2016). Development and longitudinal relationships between children's executive functions, prospective memory, and metacognition. *Cognitive Development*, *38*, 99–113. https://doi.org/10.1016/j.cogdev.2016.02.003
Wanless, S. B., McClelland, M. M., Tominey, S. L., & Acock, A. C. (2011). The influence of demographic risk factors on children's behavioral regulation in prekindergarten and kindergarten. *Early Education and Development*, *22*(3), 461–488. https://doi.org/10.1080/10409289.2011.536132
Woodcock, R. W., McGrew, K. S., & Mather, N. (2001). *Woodcock-Johnson III Tests of Achievement*. Riverside.
Woodcock, R. W., & Muñoz-Sandoval, A. F. (1993). An IRT approach to cross-language test equating and interpretation. *European Journal of Psychological Measurement*, *32*, 233–241.
Zelazo, P. D., Forston, J. L., Masten, A. S., & Carlson, S. M. (2018). Mindfulness plus reflection training: Effects on executive function in early childhood. *Frontiers in Psychology*, *9*(FEB), 1–12. https://doi.org/10.3389/fpsyg.2018.00208
Zelazo, P. D., Müller, U., Frye, D., Marcovitch, S., Argitis, G., Boseovski, J., Chiang, J. K., Hongwanishkul, D., Schuster, B. V., & Sutherland, A. (2003). The development of executive function in early childhood. *Monographs of the Society for Research in Child Development*, *68*(3), vii–137. https://doi.org/10.1111/j.0037-976x.2003.00260.x
Zimmerman, B. J., & Schunk, D. H. (2012). *Self-regulated learning and academic achievement: Theory, research, and practice*. Springer Science & Business.

How to cite this article: Gonzales, C. R., Merculief, A., McClelland, M. M., & Ghetti, S. (2021). The development of uncertainty monitoring during kindergarten: Change and longitudinal relations with executive function and vocabulary in children from low-income backgrounds. *Child Development*, 00, 1–16. https://doi.org/10.1111/cdev.13714