Children's Intuitive Theories of Academic Performance

Melis Muradoglu

Andrei Cimpian

New York University

In press, *Child Development*.

Acknowledgments

Research reported in this publication was supported by grants from the National Science Foundation (BCS-1733897), the Mindset Scholars Network, and the Raikes Foundation, awarded to Cimpian. We are grateful to the children and parents that participated in this research; to the Cognitive Development Lab for helpful comments on previous drafts of the manuscript and for their assistance with data collection; and to Zhi Xun for study illustrations.

Correspondence concerning this article should be addressed to Melis Muradoglu, Department of Psychology, New York University, 6 Washington Place, New York, NY 10003. E-mail: melis.muradoglu@nyu.edu

Abstract

How do children reason about academic performance across development? A classic view suggests children's intuitive theories in this domain undergo qualitative changes. According to this view, older children and adults consider both effort and skill as sources of performance (i.e., a "performance = effort + skill" theory), but younger children can only consider effort (i.e., a "performance = effort" theory). Results from two studies (N = 240 children aged 4–9) contradict the claim of theory change, suggesting instead that children as young as 4 operate with an intuitive theory of academic performance that incorporates both effort and skill as explanatory concepts. This work reveals that children's understanding of academic performance is more continuous across development than previously assumed.

Keywords: achievement; concepts; intuitive theories

Children's Intuitive Theories of Academic Performance

How do young children understand the sources of academic performance? What do they think are the "ingredients" that combine to lead a person to succeed or fail in school and other academic pursuits? This understanding, which is likely to take the form of an intuitive causal-explanatory theory (e.g., Carey, 1985; Gopnik, Meltzoff, & Kuhl, 1999; Keil, 1995; Wellman & Gelman, 1992), undoubtedly shapes children's experiences in school. For example, to the extent that young children's theories suggest that performance depends entirely on effort (e.g., Nicholls, 1990), they may be particularly resilient and optimistic in school because—according to their intuitive theory—trying harder guarantees doing better. In contrast, if young children's intuitive theories use other concepts besides effort to explain performance (e.g., skill), they may be susceptible to negative emotions and feelings of helplessness when they fail, much like older children are (e.g., Cimpian, Arce, Markman, & Dweck, 2007). Despite tremendous progress over the past several decades in our scientific understanding of achievement-related cognitions (e.g., Dweck, 1999, 2006; Nicholls, 1984; Senko, Hulleman, & Harackiewicz, 2011; Wigfield et al., 2015), the literature to date does not provide a clear answer to a crucial question: Is the content of children's intuitive theories of performance relatively continuous across development, or does it undergo qualitative changes? The goal of the present experiments is to reconcile some of the inconsistent findings on this topic and make progress toward determining how children understand the sources of academic performance across development.

Continuity vs. Change in Children's Intuitive Theories of Performance

To make sense of their observations, young children rely on abstract intuitive theories. Intuitive theories can be thought of as domain-specific cognitive structures that, like scientific theories, permit interpretation, explanation, and prediction (Gelman, Coley, & Gottfried, 1994;

Gopnik & Wellman, 1992).[1] At the core of an intuitive theory are the concepts that explain the phenomena that fall under the theory's scope (in our case, academic performance; e.g., Carey, 1985). (We use the terms "academic performance" and "performance" interchangeably throughout the paper. However, we acknowledge that theories of performance may differ across domains [e.g., academics vs. sports].) Most researchers agree that the intuitive theory of performance that is common among older children and adults relies on two key concepts: *effort* and *skill* (which some also term *ability* or *competence*; e.g., Dweck & Leggett, 1988; Kun & Parsons, 1974; Nicholls, 1978, 1990; Nicholls & Miller, 1984a, 1984b). We will refer to this theory as the "performance = effort + skill" theory (see Figure S1 in the Supplementary Online Materials [SOM]). According to it, a person's performance in a particular context can be explained as a function of the amount of effort they put into their performance and the level of skill they (currently) possess. Here, *effort* consists of behaviors such as concentrating and spending time on the task being performed, paying attention, avoiding distractions, and so on. These behaviors are largely under the person's control when performance is being assessed. In contrast, *skill* is a disposition and is not under the person's control at the moment when performance is assessed. A person cannot become instantly more skilled, at will; rather, they have a certain level of skill at any one point in time, and it is that particular level that factors into their performance (e.g., Weiner 1985). Note, however, that skill is at least partly malleable—and thus under one's control—over *broader* temporal scales (i.e., over weeks, months, years). The extent to which skill is seen as malleable over time differentiates what are known as growth mindsets (which assume high malleability) from fixed mindsets (which assume low malleability;

[1] We use the term *theory* to describe a set of related causal concepts or ideas (here, effort and skill) that people appeal to when making sense of events (here, events involving academic performance). Intuitive theories lack the reliability and coherence of formal scientific theories. However, both share the more basic property of consisting of causal elements that are recruited for explanation and prediction.

Dweck, 1999, 2006). But, importantly, growth and fixed mindsets both agree that, at a particular point in time, a person has a certain level of skill that influences how well they do, alongside the effort they put into their performance.

Similar to a scientific theory, the "performance = effort + skill" intuitive theory can be used to formulate predictions and draw various inferences. For example, this intuitive theory licenses the prediction that a person who put in more effort will perform better than a person who put in less effort (holding skill constant), and that a person who has more skill will perform better than a person who has less skill (holding effort constant). We can also infer a person's level of skill if we know how they performed and how much effort they put in (and vice-versa; Heider, 1958). For instance, if they did very well without putting in much effort, the "performance = effort + skill" equation suggests that they have a high level of skill. More generally, the theory licenses a general expectation of an inverse relationship between effort and skill (e.g., Amemiya & Wang, 2018; Dweck & Bempechat, 1983; Kun, 1977; Nicholls, 1990; Nicholls & Miller, 1984; but see Muenks, Miele, & Wigfield, 2016). We expect that people who are more skilled will have to put in less effort to perform at a certain level, whereas those who are less skilled will have to compensate by devoting more effort to the task.

Although older children and adults reason this way about performance, it is much less clear whether younger children also share this theory, hence the focus of the present research. One line of research has put forth what we might call a *theory-change view,* which was first proposed in the late 1970s and continues to be influential today. According to this view, children start out with an impoverished, "performance = effort" intuitive theory, which explains performance solely in terms of the effort put in. Then, at some point during middle childhood, children transition to the "performance = effort + skill" intuitive theory that we described above

(e.g., Droege & Stipek, 1993; Harari & Covington, 1981; Nicholls, 1978; Nicholls & Miller, 1983, 1984b; for reviews, see Dweck, 2002; Muenks & Miele, 2017; Nicholls & Miller, 1984a). This *theory-change view* contrasts with what we might call the *continuity view*. According to the latter view, the "performance = effort + skill" theory characterizes children's understanding from an early age (e.g., Cimpian, Hammond, Mazza, & Corry, 2017; Heyman & Compton, 2006; for reviews, see Butler, 2005; Cimpian, 2017), and the concepts that children use to understand performance (namely, effort and skill) are largely continuous from preschool on.

Evidence for the Theory-Change View. According to the theory-change view, the concept of *skill* (as a dispositional trait) is absent from children's theories of performance until they are at least 10 years old (e.g., Harari & Covington, 1981; Nicholls, 1978). While young children may seem to understand and produce words that refer to ability and skill (e.g., "good at," "smart") before this age, the theory-change view posits that children simply map these words onto the concept of effort (e.g., Butler, 2005; Cimpian, 2017; Nicholls, 1990). That is, the concept of skill is not yet differentiated from the concept of effort in young children's intuitive theories. In the absence of a differentiated concept of skill, younger children are left to interpret their own and others' performance just in terms of the level of effort that went into it (i.e., "performance = effort [or skill, which is the same thing]").

The most direct evidence for this view comes from a series of influential papers from the 1970s and 1980s (Chapman & Skinner, 1989; Miller, 1985; Nicholls, 1978, Nicholls & Miller, 1984b) that used variations of the same paradigm. Because this paradigm is relevant to the present research as well, we describe it in some detail: Participating children, who generally range from 5-year-olds to adolescents, are shown two children working side by side on schoolwork. One of the two protagonists spends the entire time working (e.g., looking at a

textbook, writing), while the other spends part of the time working and part of the time off-task (e.g., looking out the window, fidgeting). After they see this sequence of events, participating children are told that the two protagonists received the same score on their schoolwork. Participants are then asked a series of questions to assess their inferences about the protagonists' level of skill based on their (equal) performance and their (unequal) effort. Generally, younger children (before age 10) state that the child who worked the entire time was *more* skilled. This response was interpreted as showing that young children do not operate with a "performance = skill + effort" theory; if they did, they would use the fact that the two protagonists performed equally to infer that the one who put in less effort is more skilled. Instead, children assessed the protagonists' skill based on their effort, which suggests that skill is not yet differentiated from effort in their minds, as would be expected if they were operating with a "performance = effort" intuitive theory. Studies using this two-protagonist paradigm provide the most compelling and direct evidence for the theory-change view. Two other lines of work provide additional, albeit more indirect, support for the theory-change view: (1) young children's optimism in response to achievement failure, and (2) young children's difficulties with reasoning about dispositions.

First, young children seem to hold generally—and often excessively—positive views of themselves (e.g., Schuster, Ruble, & Weinert, 1998; Stipek, 1981; Thomaes, Brummelman, & Sedikides, 2017), as well as others (e.g., Boseovski & Lee, 2006; Lockhart, Chang, & Story, 2002). Young children's overly optimistic thinking has been documented in achievement contexts specifically: Several studies have suggested that young children do not show helpless responses to failure (e.g., Miller, 1985; Rholes, Blackwell, Jordan, & Waters, 1980) or lower expectations for the future in the face of failure (e.g., Parsons & Ruble, 1977). For example, rather than reducing their expectations of how they will do in the future after repeated failure on

a task, preschoolers maintain near-maximal expectations for future success (Stipek & Hoffman, 1980). This work indirectly supports the theory-change view by showing that failure is not demoralizing for young children—arguably, because their theories of performance suggest that effort is the only relevant factor, and thus doing better is just a matter of trying harder.

Second, the theory-change view is indirectly supported by evidence that children's understanding of dispositional traits seems to undergo qualitative changes from early to middle childhood. For instance, young children often fail to use an individual's trait-relevant behaviors (e.g., Anne shared her lunch) to guide their predictions about the individual's future behavior (e.g., Anne will help a neighbor tomorrow; Rholes & Ruble, 1984). Given that skills are dispositions that lead people to perform predictably across circumstances, these findings lend support to the idea that young children's intuitive theories of performance do not include the concept of skill, consistent with the claim of discontinuities in the content of children's theories of performance across development.

Evidence for the Continuity View. Newer evidence has eroded some of the support for the theory-change view, challenging in particular the claims about young children's (1) optimistic response to failure and (2) inability to understand dispositional traits.

First, several studies have documented that achievement failure can in fact cause negative reactions in children as young as 4 and 5, just as it does in older children (e.g., Butler, 1998; Cimpian et al., 2007, 2017; Cimpian, Mu & Erickson, 2012; Heyman, Dweck, & Cain, 1992; Kamins & Dweck, 1999; Smiley & Dweck, 1994). These negative, helpless reactions are common when it is made obvious to children that they have failed (e.g., as when a parent or teacher criticizes their performance; Hebert & Dweck, 1985; Heyman, Dweck, & Cain, 1992; Kamins & Dweck, 1999; Smiley & Dweck, 1994) and when the potential link between failure

and low skill is highlighted in the experimental context (e.g., via linguistic cues such as generic statements; Cimpian et al., 2007, 2012).

Second, there is now considerable evidence that children as young as 3 or 4 can reason appropriately about dispositional traits (e.g., Boseovski & Lee, 2006; Hermes, Behne, & Rakoczy, 2015; Heyman & Gelman, 1999). The early studies on children's understanding of dispositional traits underestimated children's conceptual competence, in part because these studies required children to both infer a disposition from a behavior *and* use this inferred disposition to make predictions about future behavior. Young children may have trouble with this task simply because it requires that they string two separate inferences together, not because they lack the concept of dispositional traits. Each of the separate inferences is in fact within young children's repertoire (e.g., Heyman & Gelman, 1999). In other words, even young children can use information about dispositions to guide inferences about future behavior.

This evidence suggests that, contrary to earlier claims, young children (1) are not always irrationally optimistic in the face of failure and (2) are capable of inferring and reasoning in terms of dispositional traits. Although this evidence is inconsistent with the theory-change view, the specific claims it contradicts are somewhat peripheral to this view. The findings from the two-protagonist paradigm (e.g., Nicholls, 1978), which provide the most striking evidence for the theory-change view, have gone largely unchallenged and are still generally thought to provide an accurate portrayal of the development of children's intuitive theories of performance (for a review, see Cimpian, 2017). The present studies will test an alternative interpretation of children's responses in the standard two-protagonist paradigm.

The Present Research

We hypothesize that two methodological features of the canonical two-protagonist

paradigm (Nicholls, 1978) may lead young children to respond as if they conflated skill and effort, artificially giving the impression of a transition from a "performance = effort" to a "performance = effort + skill" theory across the span from ages 5 to 10. These methodological issues, which we detail below, raise doubts about the interpretation of the results obtained with this paradigm. We predict that modifying these features—with minimal changes to the canonical version of the paradigm—will reveal evidence of the "performance = effort + skill" theory even in children as young as 4 and 5.

One potential issue with the canonical two-protagonist paradigm is that the experimenter always asks about the effort that the protagonists put in before asking about the protagonists' skill level. The fact that the experimenter's first question is about effort may signal to young children that this dimension is relevant to their answers to the subsequent question about the protagonists' skill. The task demand created by prefacing the skill question with the effort question may be particularly strong given the emphasis that is placed in young children's socialization on paying attention, following rules, and trying one's best (e.g., Eccles, Midgley, & Adler, 1984; Stipek & Daniels, 1988). To test whether this feature of the canonical two-protagonist paradigm affected children's responses, we manipulated whether the skill question was preceded by the effort question or by a question about the protagonists' mental states. Thus, half the children were randomly assigned to receive the canonical question sequence (the *effort-first condition*); for the other half, we replaced the question about effort with a question about whether the protagonists found the task difficult (the *difficulty-first condition*). Previous work has suggested that children use mental state information when drawing inferences about skill, judging someone who thought a task was easy to be smarter than someone who thought the same task was difficult (Heyman et al., 2003; Heyman & Compton, 2006). Therefore, asking about

whether the two protagonists found the test easy or difficult before asking about their skill might allow children to respond in "mature" ways by making relevant information more cognitively accessible to them. We predicted that the children who receive the difficulty question first would be more likely to respond in ways that are consistent with a "performance = effort + skill" intuitive theory (i.e., that the child who put in less effort for the same outcome is more skilled) than the children who receive the effort question first, as in the canonical paradigm.

The second methodological feature that may be problematic is an ambiguity in the sequence of events shown to children: Although the events portrayed were intended to convey that the inattentive protagonist needed less effort to get the same grade, this protagonist's behavior (that is, working on the test intermittently, while fidgeting and looking out the window) could have also signaled that the protagonist didn't know the answers, was having trouble, or disliked the task. Thus, young children may have judged the protagonist who worked intermittently as less skilled simply because they interpreted the protagonist's intermittent pattern of effort as a sign of difficulty. To test this possibility, we manipulated whether the protagonist who expended less effort on the task worked intermittently, as in the canonical task, or continuously. To clarify, the amount of effort put in by this protagonist was equally low in the two conditions; all that differed was whether their effort was spread intermittently throughout the task (the *intermittent-low-effort condition*) or concentrated at the beginning (the *continuous-low-effort condition*). We predicted that this simple change would disambiguate the sequence of events for children, while preserving the crucial contrast between the two protagonists in the amount of effort they put into the task. In other words, we predicted that children who see the low-effort protagonist working continuously (vs. intermittently) would show evidence of a "performance = effort + skill" intuitive theory from a young age, judging the protagonist that put

in less effort (for the same outcome) to be the more skilled of the two.

To summarize, the goal of the present research was to investigate whether children's intuitive theories of academic performance undergo change across childhood—initially lacking but subsequently incorporating a differentiated concept of skill. Specifically, we examined children's theories in the context of the canonical two-protagonist paradigm (e.g., Nicholls, 1978). This paradigm has provided striking evidence for the theory-change view and is, we suspect, a key reason why this view is still influential. We predicted that this paradigm would actually provide evidence for theory *continuity* when the two methodological issues discussed above are removed (Experiment 1). Because there is no debate about the intuitive theories of children older than 10, we focused on 4- to 9-year-olds in the first study and tested whether they show evidence of the more-sophisticated, "performance = effort + skill" intuitive theory. In Experiment 2, we explored this same question by further simplifying the two-protagonist paradigm and focusing more narrowly on the younger half of our age range (4- to 6-year-olds).

Experiment 1

Method

Participants. Participants were 192 children between the ages of 4 and 9 years (96 boys, 96 girls; M_{age} = 6.97 years, *SD* = 1.72 years, range = 4.04 to 9.95), recruited from a large city in the Northeastern US between October 2016 and April 2017. Participants were tested in a museum (*n* = 112), in a university laboratory (*n* = 43), or in a classroom at their school (*n* = 37). Participants were 40% White, 16% Black, 6% Asian, 7% Hispanic or Latinx, 1% American Indian or Alaskan Native, 18% Multiracial or Multiethnic, and 4% Other; 8% of parents did not report their child's race or ethnicity. The median household income was $140,000; 45% percent of parents did not report their household income. Four additional children were tested but

excluded because they refused to complete the study.

Materials. The materials presented to children in this study were modeled closely on the films shown in Nicholls (1978). Children were read storybooks that depicted two characters of the same gender (matched to children's own gender) working side by side on a test (see Tables S1 and S2 for images and script). Each storybook consisted of six pages, with each page corresponding to a temporal stage of the test (much like a frame in a video). One protagonist was depicted working continuously on the first five pages (hereafter, the *more-effort protagonist*), while the other protagonist (hereafter, the *less-effort protagonist*) was depicted working on two of the pages and doing off-task behavior (e.g., looking out the window) on three pages. On the final page of the storybook, it was revealed that both characters earned a perfect score on the test.

Design. Children were randomly assigned to one of four cells of a 2 × 2 design specified by crossing two between-subjects factors. The first factor was whether the first question children were asked (after the storybook) pertained to how hard the characters worked (effort-first condition) vs. how difficult the characters found the test (difficulty-first condition). The second factor was whether the two pages in which the less-effort protagonist was working were interleaved with pages in which (s)he was off-task (intermittent-low-effort condition) vs. shown together at the beginning of the book (continuous-low-effort condition).

Measures. After the experimenter read the two-protagonist storybook to the participant, children were asked to judge either how much effort each of the protagonists put in or how difficult each of the protagonists found the task. Which of these questions was asked first constitutes the first between-subjects factor listed above. Whichever question was not asked immediately after the vignette was asked last, after children answered the three skill questions, which are described later in this section.

The wording of the question about effort was, “Did (s)he work hard on the test or did (s)he not work hard?” If the child stated that the protagonist worked hard, the experimenter followed up with a three-point scale consisting of three circles of increasing diameter that the child could point to (“How hard did (s)he work? A little hard, hard, or really hard?”). Thus, children’s response to this question could range from 1 to 4 (1 = *did not work hard*, 2 = *worked a little hard*, 3 = *worked hard*, 4 = *worked really hard*; see Table S5 for means). The wording of the question about whether the protagonists experienced difficulty was, “Did (s)he think the test was easy to do or hard to do?” The experimenter followed up with the same three-point scale as above (“How easy/hard did (s)he think the test was? A little easy/hard, easy/hard, or really easy/hard?”). Children’s response to this question could range from 1 to 6 (1 = *really easy*, 2 = *easy*, 3 = *a little easy*, 4 = *a little hard*, 5 = *hard*, 6 = *really hard*; see Table S5 for means).

Next, children were asked three questions that assessed their concept of skill. Two of these questions were used to gauge whether the children are operating with a “performance = effort” or a “performance = effort + skill” theory. These two questions are described first below. The third question assessed children’s mindsets about the malleability of skill.

Children were first asked about which protagonist is more skilled (Nicholls, 1978): “Is one of these kids smarter, or are they the same?” If children stated that they thought that one of the protagonists was smarter, children were then asked to choose which protagonist (−1 = *more-effort protagonist is smarter*, 0 = *the protagonists are the same*, 1 = *less-effort protagonist is smarter*). Because this question provides the most direct test of children’s intuitive theories, it was always asked first (among the three questions about skill) to prevent influence from the other questions and from the extra vignette that accompanied the prediction question, described next. To clarify, although the term “smart” can also refer to raw intellectual capacity (rather than

skill per se), when children are asked to judge someone based on the outcome of a test in school, the most likely interpretation is that "smart" refers to competence and skill (in the same way that people use "I'm smart at this" interchangeably with "I'm good at this"). We should also note that this is the terminology used in the canonical two-protagonist paradigm (see Nicholls & Miller, 1984b). The other question that assessed children's intuitive theories of performance asked them to predict the two protagonists' performance under conditions of equal effort (also adapted from Nicholls, 1978). Rather than presenting this question verbally, as in Nicholls (1978), we scaffolded children's understanding with visuals—the two protagonists working side by side in five consecutive pages of a (new) storybook—before asking them for their prediction about who got a higher score (−1 = *more-effort protagonist* or 1 = *less-effort protagonist*; see Table S3). Responses to these two key questions (evaluation and prediction) were scaled from −1 to 1, such that −1 indicates a "performance = effort" theory and 1 indicates a "performance = effort + skill" theory. The neutral midpoint of this scale was 0; thus, responses close to 0 indicated that children did not endorse one theory over the other.

The remaining (third) question measured children's beliefs about the malleability of skill or ability (i.e., their mindsets; Dweck, 2006): "Can anyone be really smart at school if they wanted to be? Or are there some kids who could never be really smart at school?" (adapted from Heyman & Compton, 2006). This question was followed with a confidence rating: "Are you sort of sure? Or really sure?" Responses were scaled from −1 to 1, with −1 indicating a growth mindset and 1 indicating a fixed mindset (−1 = *anyone* and *really sure*, −.33 = *anyone* and *sort of sure*, .33 = *not anyone* and *sort of sure*, 1 = *not anyone* and *really sure*). This mindset question was included to test the discriminant validity of our manipulations. Disambiguating the features of the canonical two-participant paradigm should allow children to reveal that they possess a

concept of skill (distinct from effort) but should not change whether they think skill is fixed or malleable. The order of the prediction question and the mindset question was counterbalanced.

To reiterate, the evaluation and prediction questions were intended to tap children's intuitive theories of performance. Consistent with their intended function, these two questions were significantly correlated, $r = .27$, $p < .001$. In both of the studies reported here, these two questions are analyzed together and are jointly referred to as "children's skill judgments." The mindset question correlated neither with the evaluation question, $r = -.06$, $p = .44$, nor with the prediction question, $r = .04$, $p = .62$. These near-zero correlations confirm that the mindset question assesses a different aspect of children's achievement cognitions than the other two questions. Thus, we analyzed the mindset question separately from the other two.

Data Analysis. We submitted children's responses to the evaluation and prediction questions to a multilevel mixed-effects linear model with the content of the first question (effort-first = 0 vs. difficulty-first = 1; dichotomous level-2 predictor), the type of low effort (intermittent-low-effort = 0 vs. continuous-low-effort = 1; dichotomous level-2 predictor), children's age (continuous level-2 predictor), and all possible interactions as predictors. We also included crossed random intercepts for subject and question in the model. This model was computed using the *mixed* command in Stata 15.1. We analyzed children's responses to the mindset question with a linear regression model using the predictors above. This model was computed using the *regress* command in Stata 15.1. Cohen's *d* was computed from observed means as a measure of effect size. Preliminary analyses suggested that children's gender was unrelated to their responses and did not moderate any effects reported below, so we did not consider it further. We mean-centered all predictors to facilitate interpretation of coefficients.

Open Data and Analytic Syntax

The raw data for both experiments and the analytic syntax are openly available on the Open Science Framework (OSF): https://osf.io/2g36q/?view_only=6eca8c1f8bbc42a59d287cf6ffab2754. This repository has been anonymized for review.

Results and Discussion

We predicted that children would respond in line with a "performance = effort + skill" theory when they are (1) not asked to judge the protagonists' effort before judging their skill (which could create a task demand) and (2) not shown the less-effort protagonist working intermittently (which could be misinterpreted as experiencing difficulty). If both of these problematic features of the canonical two-protagonist paradigm are removed, we expect that even young children would judge the protagonist who put in less effort but achieved the same outcome as being more skilled; similarly, children should predict that this protagonist will do better when effort is equated.

Consistent with these predictions, children were significantly more likely to judge the protagonist who put in less effort as smarter (1) when first asked about whether the protagonists found the test easy or difficult ($M = -0.05$, $SD = 0.70$) than when first asked about whether the protagonists worked hard on the test ($M = -0.41$, $SD = 0.59$), $b = 0.36$, $SE = 0.09$, $z = 4.14$, $p < .001$, $d = 0.55$, and (2) when the protagonist who put in less effort worked continuously at the beginning of the test ($M = -0.03$, $SD = 0.71$) rather than intermittently at various points during the test ($M = -0.44$, $SD = 0.56$), $b = 0.41$, $SE = 0.09$, $z = 4.75$, $p < .001$, $d = 0.64$. (In Appendix S1, we detail evidence suggesting that the effect of the continuous- vs. intermittent-low-effort manipulation was mediated by the perception that the protagonist in the intermittent-low-effort condition found the test harder.) We also found a significant interaction between the two factors, $b = 0.39$, $SE = 0.17$, $z = 2.24$, $p = .025$ (see Table S4 for the full model output). This interaction

was driven by the fact that the scores in one of the four design cells—specifically, the cell where both problematic features of the canonical paradigm were removed—were substantially higher than the scores in the other three cells, which were similarly low (see Figure 1; the separate results for the evaluation and prediction questions are displayed in Figures S2 and S3). Children's age did not moderate either the two main effects or the interaction, and there was also no overall relationship between age and skill judgments (all *p*s > .247).

Next, we tested whether children showed systematic endorsement of the "performance = effort" theory or the "performance = effort + skill" theory in each of the four cells of our 2 × 2 design (see Figures 1 and 2). If the average in a cell is significantly above 0 (the neutral midpoint), it suggests systematic endorsement of a "performance = effort + skill" theory, while an average that is significantly below 0 reflects a "performance = effort" theory. Children's scores were most negative in the cell that most closely resembles the canonical two-protagonist paradigm (effort-first and intermittent-low-effort), $M = -0.52$, $SD = 0.54$, $t(47) = -6.74$, $p < .001$, $d = -0.97$.[2] (This replicates prior findings with the canonical paradigm [e.g., Nicholls, 1978], which suggested that most 4- to 9-year-olds do not consider skill to be a distinct explanatory concept.) In contrast, the cell in which these two features were removed (difficulty-first and continuous-low-effort) was the only one with scores significantly above 0, $M = 0.25$, $SD = 0.68$, $t(47) = 2.53$, $p = .015$, $d = 0.37$. In this cell, children judged the low-effort protagonist to be more skilled and predicted that (s)he would do better when effort is equated across the two protagonists. Scores in the remaining two cells of the design, where only one problematic feature was removed, were significantly below 0 (*p*s < .002)

[2] For comparison, we recruited a sample of adult participants ($N = 23$; $M_{age} = 19.22$ years) to whom we administered an online version of this condition. As expected, adults' responses ($M = 0.26$, $SD = 0.69$) were significantly more in line with a "performance = effort + skill" theory than children's responses ($M = -0.52$, $SD = 0.54$), $t(69) = 5.24$, $p < .001$.

Our final analysis focused exclusively on the design cell that, in our view, provides the most apt test of children's theories of performance (i.e., the difficulty-first and continuous-low-effort cell). Our goal in this analysis was to estimate with more precision the age at which children show evidence of a "performance = effort + skill" theory (i.e., their scores are significantly above 0; see Figure 2). Inspired by the Johnson-Neyman "regions of significance" approach (Johnson & Neyman, 1936; see Preacher, Curran, & Bauer, 2006), we used the mixed-effects model described above to estimate marginal means (and their standard errors) at 0.1 year increments between ages 4 and 10. This analysis suggested that children's scores in this cell of the design were significantly above 0 starting at 5.6 years. That is, contrary to claims of discontinuity in the theories children use to understand academic performance, even 5-year-olds seem to understand that performance is a function of both effort and skill.

No Effects for the Mindset Question. The regression on children's responses to the mindset question revealed no main effects or interactions involving the factors we manipulated (*p*s > .180; see Table S6 and Figure S4), suggesting that these experimental manipulations did not cause shifts in children's beliefs about whether skill is fixed or malleable. This result provides evidence for the discriminant validity of our manipulations and confirms that mindsets are orthogonal to the more-basic theories that we are investigating here, which are neutral with respect to the malleability of one's skill over the long term (see Figure S1). Children did, however, endorse the idea that anyone could be smart more with age, $b = -0.08$, $SE = 0.03$, $t = -2.83$, $p = .005$, showing evidence of a stronger growth mindset (see Figure S5).

Conclusion

By simply disambiguating two features of the canonical paradigm used to test children's intuitive theories of performance, Experiment 1 provided evidence that children as young as 5

years of age operate with a "performance = effort + skill" theory. Thus, in contrast to prominent views on this topic, even young children consider skill to be a source of academic performance.

Experiment 2

In Experiment 1, the youngest age at which we observed systematic endorsement of a "performance = effort + skill" intuitive theory was 5.6 years. One goal of Experiment 2 was to explore whether we might find endorsement of this theory in even younger children if we further simplified the canonical two-protagonist paradigm. To this end, we introduced two modifications to the paradigm. First, because reasoning about multiple protagonists at the same time might tax young children's cognitive resources, in Experiment 2 we presented and asked children to evaluate the two protagonists *sequentially* rather than simultaneously. Second, we amplified the contrast between the effort put in by the two protagonists. In Experiment 2, the less-effort protagonist was said to finish before *an entire class* of students rather than before a single student; analogously, the more-effort protagonist was said to finish *after* an entire class of students (see Heyman & Compton, 2006). To further highlight these differences in effort, we emphasized to children that the less-effort protagonist needed less time than other students, whereas the more-effort protagonist needed more time. We predicted that this simplified two-protagonist paradigm would make it easier for young children to reveal their understanding that a person who can achieve the same outcome as another with less effort is probably more skilled, which would suggest that their intuitive theories of performance incorporate a concept of skill that is differentiated from the concept of effort. Given this experiment's goal of exploring theory continuity vs. change in early development, the participants consisted only of 4- to 6-year-olds.

In addition to exploring whether the conclusions of Experiment 1 extend to younger children, Experiment 2 tested their robustness. Recall that children in Experiment 1 responded in

accord with a "performance = effort + skill" theory only when first asked to evaluate whether the two protagonists found the test easy or difficult. In Experiment 2, we tested whether this preparatory question is necessary. That is, would young children judge the protagonist who put in less effort for the same outcome to be more skilled even if they weren't first prompted to reason about the protagonists' psychological states? To investigate this, we omitted the difficulty question altogether and asked children about the protagonists' skill without any preamble.

Method

Participants. Participants were 48 4- to 6-year-olds (24 boys, 24 girls; M_{age} = 5.55 years, *SD* = 0.85 years, range = 4.02 to 6.90) recruited from a large city in the Northeastern US during the summer of 2017. Participants were tested either in a museum (*n* = 41) or in a university laboratory (*n* = 7). Participants were 38% White, 8% Black, 6% Asian, 13% Hispanic or Latinx, 17% Multiracial or Multiethnic, and 10% Other; 8% of parents did not report their child's race or ethnicity. The median household income was $150,000; 40% of parents did not report their household income. Eight additional children were tested but excluded because they refused to complete the study (*n* = 4) or failed more than half of the attention checks (*n* = 4; see below).

Materials. Each child heard about two protagonists (gender-matched to the child), presented sequentially. Both protagonists were said to have received a "really high score" on a test, but one needed "very little time," finishing before everyone else in his or her class, whereas the other needed "a very long time," finishing after everyone else in his or her class (for the full script, see Table S7). The presentation order of the protagonists was counterbalanced across participants. As in Experiment 1, children saw a schematic drawing of each protagonist that depicted a young child of the relevant gender. Children were asked two attention check questions per vignette (for a total of four) to ensure that they encoded the key information about the

protagonists' performance and whether they finished the test quickly or slowly. Children who answered three or more of these questions incorrectly were excluded (n = 4).

Measures. After the attention check questions, children were asked two test questions (in counterbalanced order across participants). Both questions were about the protagonists' skill and spoke to whether young children differentiate skill from effort (in which case they should judge that the less-effort protagonist is more skilled) or not (in which case they should judge that the more-effort protagonist is more skilled). Children's responses to these questions were scaled from 0 to 1, with higher scores reflecting greater perceived skill.

The first question asked children to evaluate the relevant protagonist's skill: e.g., "How smart is Jamie? Is she smart or not very smart?" If children stated that the protagonist was smart, the experimenter followed up with a three-point scale accompanied by a schematic drawing of circles of increasing diameter: "Is she sort of smart, smart, or really smart?" (0 = *not very smart*, .33 = *sort of smart*, .67 = *smart,* 1 = *really smart*). The second question asked children to make a prediction about the relevant protagonist's future performance: e.g., "On another day at school, Jamie was playing a game for kids that are really, really smart. Only kids who are really smart can be good at this game. What kind of score do you think Jamie got in this game? Did she get a good score or a not-so-good score?" The experimenter followed up on children's response to this initial question with a three-point scale accompanied by three schematic faces with smiles or frowns of increasing intensity: "Was her score sort of [not] good, [not] good, or really [not] good?" (0 = *really not good*, .2 = *not good*, .4 = *sort of not good*, .6 = *sort of good*, .8 = *good*, 1 = *really good*). Before the experimenter asked whichever of the test questions came second, she reminded the child of the protagonist's performance ("a really high score") and whether they needed more vs. less time than other children. The evaluation and prediction questions were

significantly correlated (r = .48, p < .001) and were analyzed together in a single model.

Data Analysis. We submitted children's responses to the evaluation and prediction questions to a multilevel mixed-effects linear model with protagonist (less-effort protagonist = 0 vs. more-effort protagonist = 1; dichotomous level-1 predictor), children's age (continuous level-2 predictor), and their interaction as predictors. Preliminary analyses revealed substantial order effects, so we also included protagonist order (more-effort protagonist is first = 0 vs. less-effort protagonist is first = 1; dichotomous level-2 predictor) and its interactions with the other two predictors in the model. The model included crossed random intercepts for subject and question.

Results and Discussion

If 4- to 6-year-olds operate with a "performance = effort + skill" intuitive theory, we would expect them to judge the protagonist who got a good grade with less effort to be more skilled. In contrast, if children operate with an intuitive theory that does not differentiate skill from effort ("performance = effort"), we would expect them to show a difference in the opposite direction, rating the protagonist who put in more effort as more skilled.

Children judged the less-effort protagonist (M = 0.84, SD = 0.25) as more skilled than the more-effort protagonist (M = 0.72, SD = 0.32), b = −0.12, SE = 0.04, z = −3.28, p = .001, d = −0.39, indicating the presence of a differentiated concept of skill in this young group (see Table S8 for the full model). In addition, this difference was not moderated by children's age, b = −0.06, SE = 0.04, z = −1.43, p = .15. However, the perceived skill difference between the two protagonists was qualified by a significant interaction with protagonist order, b = −0.21, SE = 0.07, z = −2.77, p = .006. Unpacking this interaction, we found that the difference between the two protagonists emerged only when children judged the less-effort protagonist first (see Figures 3 and 4). This interaction is likely due to a ceiling effect in the other protagonist order: Children

who were presented with the more-effort protagonist first rated him or her as highly skilled (M = 0.79, SD = 0.26), which is reasonable given that the protagonist was said to have obtained a really high score; however, children had little room to subsequently increase their skill ratings when told about the less-effort protagonist (M = 0.81, SD = 0.25), $b = -0.02$, SE = 0.05, $z = -0.36$, $p = .72$. Although children who received the less-effort protagonist first also rated him or her as highly skilled (M = 0.86, SD = 0.24), they then had substantially more room to *decrease* their rating when judging the more-effort protagonist (M = 0.65, SD = 0.35), $b = -0.22$, SE = 0.05, $z = -4.27$, $p < .001$ (see Figures S6 and S7 for separate graphs for the evaluation and prediction questions).

We next sought to determine with more precision the age at which children in our sample showed systematic evidence of the "performance = effort + skill" theory, as indicated by significantly higher skill ratings for the less-effort than the more-effort protagonist. Because of the ceiling effect observed when the more-effort protagonist was presented first, in this analysis we focused just on the children for whom the less-effort protagonist was first. The prediction was that the changes we made in Experiment 2 to simplify the standard two-protagonist paradigm might allow even children younger than 5.6 years to show evidence of the "performance = effort + skill" theory. This analysis indicated that children rated the less-effort protagonist as significantly more skilled than the more-effort protagonist starting at 4.8 years (see Figure 4). Thus, Experiment 2 provides evidence of a differentiated concept of skill in 4-year-olds, approximately 10 months earlier than in Experiment 1. Notably, the fact that the skill questions in this study were not preceded by a question about whether the protagonists experienced difficulty arguably makes the present setup a more conservative test of children's theories.

General Discussion

The present studies investigated the development of children's intuitive theories of academic performance, focusing in particular on whether young children consider skill, in addition to effort, to be a source of achievement outcomes. Although the past few decades have brought about considerable progress in the field's understanding of achievement beliefs, the question of continuity vs. change in children's intuitive theories of performance remains open. According to the *theory-change view*, older children and adults consider both effort and skill as sources of performance, but children younger than 10 or 11 only consider effort. The present studies provide support for the *continuity view* by demonstrating that children as young as 4 have intuitive theories of performance that include skill as a distinct causal-explanatory concept.

Historically, the theory-change side of the debate has been dominant (e.g., Harari & Covington, 1981; Kun, 1977; Nicholls, 1978). Several classic lines of research provided support for the idea that young children's thinking about performance is qualitatively different from that of older children (e.g., their explicit judgments of skill seem to treat it as interchangeable with effort; they seem irrationally optimistic in the face of failure or criticism). More recently, however, critiques of some of these classic findings have begun to surface (for a review, see Cimpian, 2017), which has provided some momentum for the continuity side of the debate. To advance this debate, our experiments tested an alternative explanation for the findings that currently provide the strongest, most direct support for the theory-change view—namely, the findings on children's explicit judgments of ability and skill in the paradigm where two protagonists put in different amounts of effort and achieve the same outcome. In this paradigm, young children seem to judge individuals who need to exert considerable effort for a certain outcome to be *more* skilled than individuals who can achieve the same outcome with less effort (Nicholls, 1978; see also Droege & Stipek, 1993; Kun, 1977; Nicholls & Miller, 1984b).

We suggested that young children's answers in the canonical version of this paradigm may be partly due to task demands (e.g., always asking about the protagonists' effort first) and ambiguities in the stimuli (e.g., the intermittent working pattern of one of the protagonists may have signaled difficulty with the test). Consistent with this alternative explanation, children in Experiment 1 judged the protagonist who worked harder to also be smarter *only* when one or both of these features were present (see Figure 1). When both features were absent, the 4- to 9-year-olds in this study judged the protagonist who put in less work for the same outcome to be more skilled, consistent with a "performance = effort + skill" theory. The results obtained with this revised version of the two-protagonist paradigm suggested that children as young as 5.6 years have access to, and reason with, a differentiated concept of skill. In Experiment 2, we further clarified and simplified the two-protagonist paradigm (e.g., by presenting the protagonists sequentially) and found evidence of this concept even in children below the age of 5. Even 4-year-olds—especially above the age of 4.8 years—systematically inferred that an individual who finished a task quickly (i.e., needed little effort) had more skill than an individual who, though equally achieving, took longer to finish (i.e., needed more effort). Children made these inferences even though they had not been "prepared" for them with a question about whether the protagonists experienced difficulty (as in Experiment 1).

However, it is also noteworthy that in Experiment 2, evidence of these sophisticated inferences was found in only one of the two possible orders of the protagonists. When the high-effort protagonist was first, children's responses did not distinguish between the two protagonists, likely because their estimates of the protagonists' skill were already close to maximum for the first (high-effort) protagonist. These positive evaluations are in some ways reasonable, given that the high-effort protagonist was said to have earned a "really high score."

Had the outcome information been more middling, it is possible that children would have differentiated the two protagonists regardless of the order in which they were introduced. However, this result may also suggest that children did not *spontaneously* rely on information about relative effort when making their evaluations. If they had used this information spontaneously, they might have been more likely to downgrade their evaluation of the high-effort protagonist even when they saw this protagonist first. Rather, children may have been prompted to use the effort information by the contrast between the amounts of effort put in by the two protagonists in our task. Even so, the fact that children responded systematically to this contrast suggests that they understood that a person who has to expend more effort to achieve the same outcome is less skilled—that is, these results suggest that young children's theories of performance incorporate both effort and skill as explanatory constructs.

It is important to acknowledge that other recent work has also pointed to greater-than-expected sophistication in children's reasoning about competence. For instance, like older children and adults, younger children consider objective performance (e.g., how many problems someone got right) and perceptions of task difficulty (e.g., whether someone found the test easy or difficult) when making inferences about skill (Heyman et al., 2003; Heyman & Compton, 2006; for a review, see Cimpian, 2017). Nevertheless, the claim that the concepts of effort and skill are conflated in young children's theories of performance has gone uncontested, largely because of the results from the standard two-protagonist paradigm. Since the findings from this paradigm were the main source of support for the view that intuitive theories of competence and achievement are discontinuous across development, the present work adds an important piece of evidence for the alternative view in this debate—that young children also have access to the same concepts that older children use to make sense of achievement outcomes (e.g., Butler,

1989; Cimpian et al., 2007, 2012, 2017; Heyman & Compton, 2006; Yang & Frye, 2016).

To clarify, we are not claiming that children's reasoning and behavior in achievement settings is the same in all respects across development. Achievement-related behaviors do change with age. For example, it is clear that young children are less focused on social comparisons (e.g., Ruble, Boggiano, Feldman, & Loebl, 1980) and on average more optimistic (e.g., Parsons & Ruble, 1977) than older children and adults. Along the same lines, young children do tend to demonstrate hardier reactions to failure in achievement contexts than older children and adults (e.g., Miller, 1985; Rholes et al., 1980). If differences in underlying conceptual competence are not responsible for these age-related changes, what is? One plausible explanation appeals to the systematic changes that occur in children's schooling environments (e.g., Eccles et al., 1984), which may highlight certain ways of thinking about achievement over others (e.g., Stipek & Daniels, 1988). In early childhood, most of the input children receive about their performance is centered on putting forth effort, and their performance in school and elsewhere is almost never formally evaluated. Further, early environments rarely facilitate social comparison and instead typically focus on individual progress (e.g., Rosenholtz & Rosenholtz, 1981). Thus, children's environments are likely to shape the ways of reasoning about achievement that are "default," or most accessible, at a certain point in development, even if the conceptual structures are in principle available that could lead to a different pattern of reasoning and behavior. While young children sometimes behave as if effort is the only thing that matters for performance, this should not be mistaken for conceptual immaturity—they are perfectly capable of understanding that performance is also influenced by skill, as the present data suggest.

Although this was not our main focus, it is worth noting that the present studies also reinforce the claim that young children understand behavior in terms of dispositions. The

children in our studies not only evaluated the two protagonists' skill differently, but they also used these judgments of skill to predict how the protagonists would behave in the future (see Figures S3 and S7), which suggests that children understood skill as a disposition that leads people to behave in consistent ways (e.g., perform well) across relevant situations (e.g., when taking tests). Thus, by revealing that by the age of 4 children believe that performance stems in part from a disposition (skill or ability), the present results contribute to the literature suggesting the presence of rich trait inferences in preschool-age children (e.g., Boseovski & Lee, 2006; Hermes et al., 2015; Heyman & Gelman, 1999).

Beyond its theoretical contributions, the present work also connects with issues of practical relevance for children's actual achievement in the classroom. If children as young as 4 or 5 understand that skill is an important factor that shapes academic performance, they can also begin to develop systematic beliefs about the *nature* of one's skill—that is, mindsets (e.g., Dweck, 2006). In older children and adults, mindsets influence what types of goals students adopt (e.g., learning new things vs. getting good grades), whether they show a resilient or helpless response to difficulties, as well as how they actually perform in school (for reviews, see Dweck, 1999, 2006). Due in part to the long-lasting appeal of the view that young children conflate skill with effort, there is little research on mindsets and their effects among young children—after all, if children do not understand what skill is, there is no point in investigating their ideas about its sources and malleability. The present research suggests that this stance toward mindsets in early childhood is misguided and underscores the need for research on the content and effects of mindsets in young children, as well as for reliable and valid measures of mindsets in this age group. In fact, contrary to the idea that young children overestimate the role of effort, in Experiment 1 younger children were *more* likely than older children to endorse the

idea that "some kids could never be really smart at school," which suggests a fixed mindset. This finding highlights the need for future research on mindsets in early childhood.

Limitations and Future Directions

A limitation of the present work concerns the way in which we operationalized the notion of effort. In both studies, we conveyed effort information through the amount of time the protagonists spent on the relevant task. However, effort takes on a number of other forms as well (e.g., paying attention, practicing). Although research on how children conceive of effort is limited, some evidence suggests that children's concepts in this domain encompass a wide range of behaviors, from "always working" to "staying out of trouble" (Blumenfeld, Pintrich, & Hamilton, 1986). What falls under "effort" seems to also change across development. For instance, one study found that younger children tend to think of effort in terms of conduct (e.g., following rules and listening to the teacher), while older children are instead more likely to reference persistence and a willingness to try (Lee, Carter, & Xiang, 1995). In future research, it would be useful to vary the type of effort described to participants in order to test the generalizability of the conclusion that even young children view effort and skill as distinct.

While on the topic of effort, we should note that we illustrated protagonists' low effort by portraying them as completing the task *quickly*. As a result, it is possible that children's judgments of skill were influenced by the perception that the less-effort protagonist worked not only very little but also with speed. Children sometimes rely on the speed with which someone completes a task as a source of information about whether they are skilled (Blumenfeld et al., 1986). In future research, varying the type of effort information provided to children will also help determine whether speed information is necessary for young children to judge a character who achieves the same outcome as another with less effort as being more skilled.

Future work might also broaden the assessment of children's skill judgments. In the present studies, we relied on two items: one asking children to evaluate the protagonists' skill directly, and another eliciting a prediction about the protagonists' future performance. However, the item that asked children to evaluate the protagonists' skill relied on the term "smart"; because this term can also refer to intellectual capacity more generally (rather than just to a specific skill or ability), its use here is a limitation. We opted to use "smart" to refer to skill for consistency with the canonical two-protagonist paradigm (e.g., Nicholls & Miller, 1984b). In addition, we expected that "smart" would be interpreted in a more concrete, domain-specific sense when used to evaluate someone in light of their performance on a test (i.e., as meaning smart *at something*; see Nicholls & Miller, 1984b). The fact that the results across Experiments 1 and 2 looked very similar when analyzing the evaluation ("smart") question and the prediction question separately (see Figures S2, S3, S6, and S7) is consistent with the idea that children understood "smart" to denote skill in the particular domain being tested. However, in future work it would be informative to use other child-friendly terms to refer to skill (e.g., "Who is better at this?").

Finally, the demographic composition of our samples is a limitation worth pointing out. Although children in our studies were racially and ethnically diverse, the majority of children came from middle- and upper-class backgrounds, as evidenced by the high median household income across both studies. To explore whether our findings differed for children from low- vs. high-income families, we conducted an additional set of analyses in which we added a measure of income (above vs. below the median) as a moderator in the main models for Experiments 1 and 2. Household income did not significantly moderate the main effects or interactions reported in Experiments 1 or 2 ($ps > .29$), suggesting that children's judgments of skill did not vary as a function of family income. Of course, given the restricted range of income in these samples,

more evidence will be needed to draw firm conclusions about whether this factor does or does not shape the development of children's achievement-related cognitions.

Conclusion

The present studies suggest that, contrary to prior claims, young children operate with an intuitive theory of performance that includes not just the concept of effort but also a skill component that is differentiated from effort. This work adds to a growing body of literature that suggests young children's representational capacities are richer than once assumed, and it also has practical implications for how to best motivate young children in school.

References

Amemiya, J., & Wang, M. (2018). Why effort praise can backfire in adolescence. *Child Development Perspectives, 12,* 199-203. https://doi.org/10.1111/cdep.12284

Blumenfeld, P. C., Pintrich, P. R., & Hamilton, V. L. (1986). Children's concepts of ability, effort, and conduct. *American Educational Research Journal, 23*, 95-104.

Boseovski, J. J., & Lee, K. (2006). Children's use of frequency information for trait categorization and behavioral prediction. *Developmental Psychology, 42*, 500-513. doi:10.1037/0012-1649.42.3.500

Butler, R. (1989). Mastery versus ability appraisal: A developmental study of children's observations of peers' work. *Child Development, 60,* 1350-1361. DOI: 10.2307/1130926

Butler, R. (1998). Age trends in the use of social and temporal comparison for self-evaluation: Examination of a novel developmental hypothesis. *Child Development, 69,* 1054-1073. https://doi.org/10.1111/j.1467-8624.1998.tb06160.x

Butler, R. (2005). Competence assessment, competence, and motivation between early and middle childhood. In A. J. Elliott & C. S. Dweck (Eds.), *Handbook of competence and motivation* (pp. 202–221). New York: Guilford Press.

Carey, S. (1985). *Conceptual change in childhood*. Cambridge, MA: MIT Press.

Chapman, M., & Skinner, E. A. (1989). Children's agency beliefs, cognitive performance, and conceptions of effort and ability: Individual and developmental differences. *Child Development, 60,* 1229-1238. DOI: 10.2307/1130796

Cimpian, A. (2017). Early reasoning about competence is not irrationally optimistic, nor does it stem from inadequate cognitive representations. In A. J. Elliott & C. S. Dweck (Eds.),

Handbook of competence and motivation: Theory and application (2nd ed., pp. 387-407). New York: Guilford Press.

Cimpian, A., Arce, H. M. C., Markman, E. M., & Dweck, C. S. (2007). Subtle linguistic cues affect children's motivation. *Psychological Science*, *18*, 314-316. https://doi.org/10.1111/j.1467-9280.2007.01896.x

Cimpian, A., Hammond, M. D., Mazza, G., & Corry, G. (2017). Young children's self-concepts include representations of abstract traits and the global self. *Child Development, 88*, 1786-1798. https://doi.org/10.1111/cdev.12925

Cimpian, A., Mu, Y., & Erickson, L. C. (2012). Who is good at this game? Linking an activity to a social category undermines children's achievement. *Psychological Science, 23*, 533-541. https://doi.org/10.1177/0956797611429803

Droege, K. L., & Stipek, D. J. (1993). Children's use of dispositions to predict classmates' behavior. *Developmental Psychology*, *29*, 646-656. doi:10.1037/0012-1649.29.4.646

Dweck, C.S. (1999). *Self-theories: Their role in motivation, personality, and development*. Philadelphia, PA: Psychology Press/Taylor & Francis.

Dweck, C. S. (2002). The development of ability conceptions. *Development of Achievement Motivation*, *17*, 57-88. https://doi.org/10.1016/B978-012750053-9/50005-X

Dweck, C. S. (2006). *Mindset: The new psychology of success*. New York: Random House.

Dweck, C. S., & Bempechat, J. (1983). Theories of intelligence and achievement motivation. *Learning and motivation in the classroom* (pp. 239-256). Hillsdale, NJ: Erlbaum.

Dweck, C. S., & Leggett, E. L. (1988). A social-cognitive approach to motivation and personality. *Psychological Review, 95*, 256–273. doi:10.1037/0033-295X.95.2.256

Eccles, J. S., Midgley, C., & Adler, T. (1984). Grade-related changes in the school environment: Effects on achievement motivation. In J. Nicholls (Ed.), *Advances in motivation and achievement*: *The development of achievement motivation* (pp. 283-331). Greenwich, CT: JAI Press.

Gelman, S.A., Coley, J.D., Gottfried, G., 1994. Essentialist beliefs in children: The acquisition of concept and theories. In: Hirschfeld, L.A., Gelman, S. (Eds.), *Mapping the Mind: Domain Specificity in Cognition and Culture*. Cambridge University Press, New York, pp. 341–365.

Gopnik, A., Meltzoff, A. N., & Kuhl, P. K. (1999). *The scientist in the crib: Minds, brains, and how children learn*. William Morrow & Co.

Gopnik, A., & Wellman, H. M. (1992). Why the child's theory of mind really is a theory. *Mind & Language*, *7*, 145–171. https://doi.org/10.1111/j.1468-0017.1992.tb00202.x

Harari, O., & Covington, M. (1981). Reactions to achievement behavior from a teacher and student perspective: A developmental analysis. *American Educational Research Journal, 18*, 15–28. https://doi.org/10.3102/00028312018001015

Hayes, A. F. (2013). *Introduction to mediation, moderation, and conditional process analysis: A regression-based approach.* New York: Guilford Press.

Hebert, C. A., & Dweck, C. S. (1985). Mediators of persistence in preschoolers: Implications for development. Unpublished manuscript, Harvard University, Cambridge, MA.

Heider, F. (1958). *The psychology of interpersonal relations.* New York, NY: Wiley.

Heyman, G. D., & Compton, B. J. (2006). Context sensitivity in children's reasoning about ability across the elementary school years. *Developmental Science*, *9*, 616-627. https://doi.org/10.1111/j.1467-7687.2006.00540.x

Heyman, G. D., Gee, C. L., & Giles, J. W. (2003). Preschool children's reasoning about ability. *Child Development, 74,* 516-534. https://doi.org/10.1111/1467-8624.7402013

Heyman, G. D., Dweck, C. S., & Cain, K. M. (1992). Young children's vulnerability to self-blame and helplessness: Relationship to beliefs about goodness. *Child Development, 63,* 401-415. https://doi.org/10.1111/j.1467-8624.1992.tb01636.x

Heyman, G. D., & Gelman, S. A. (1999). The use of trait labels in making psychological inferences. *Child Development, 70,* 604-619. https://doi.org/10.1111/1467-8624.00044

Hermes, J., Behne, T., & Rakoczy, H. (2015). The role of trait reasoning in young children's selective trust. *Developmental Psychology, 51,* 1574-1587. http://dx.doi.org/10.1037/dev0000042

Johnson, P. O., & Neyman, J. (1936). Tests of certain linear hypotheses and their applications to some educational problems. *Statistical Research Memoirs,* 1, 57–93.

Kamins, M. L., & Dweck, C. S. (1999). Person versus process praise and criticism: Implications for contingent self-worth and coping. *Developmental Psychology, 35,* 835-847. doi:10.1037/0012-1649.35.3.835

Keil, F. C. (1995). The growth of causal understanding of natural kinds. In D. Sperber, D. Premack, & A. J. Premack (Eds.), *Causal cognition: A multidisciplinary debate* (pp. 234–262). Oxford: Oxford University Press.

Kun, A. (1977). Development of the magnitude-covariation and compensation schemata in ability and effort attributions of performance. *Child Development, 48,* 862-873. http://dx.doi.org/10.2307/1128335

Kun, A., & Parsons, J. E. (1974). Development of integration processes using ability and effort information to predict outcome. *Developmental Psychology, 10,* 721-732. http://dx.doi.org/10.1037/h0037014

Lee, A. M., Carter, J. A., & Xiang, P. (1995). Children's conceptions of ability in physical education. *Journal of Teaching in Physical Education, 14,* 384-393. https://doi.org/10.1123/jtpe.14.4.384

Lockhart, K. L., Chang, B., & Story, T. (2002). Young children's beliefs about the stability of traits: Protective optimism?. *Child Development, 73,* 1408-1430. https://doi.org/10.1111/1467-8624.00480

Miller, A. (1985). A developmental study of the cognitive basis of performance impairment after failure. *Journal of Personality and Social Psychology, 49,* 529-538. doi:10.1037/0022-3514.49.2.529

Muenks, K., & Miele, D. B. (2017). Students' thinking about effort and ability: The role of developmental, contextual, and individual difference factors. *Review of Educational Research, 87,* 707-735. https://doi.org/10.3102/0034654316689328

Muenks, K., Miele, D. B., & Wigfield, A. (2016). How students' perceptions of the source of effort influence their ability evaluations of other students. *Journal of Educational Psychology, 108,* 438-454. http://dx.doi.org/10.1037/edu0000068

Nicholls, J. G. (1978). The development of the concepts of effort and ability, perception of academic attainment, and the understanding that difficult tasks require more ability. *Child Development, 49,* 800-814. DOI: 10.2307/1128250

Nicholls, J. G. (1984). Achievement motivation: Conceptions of ability, subjective experience, task choice, and performance. *Psychological Review, 91,* 328-346. http://dx.doi.org/10.1037/0033-295X.91.3.328

Nicholls, J. G. (1990). What is ability and why are we mindful of it? A developmental perspective. In R. J. Sternberg & J. Kolligian, Jr. (Eds.), *Competence considered* (pp. 11–45). New Haven, CT: Yale University Press.

Nicholls, J. G., & Miller, A. T. (1983). The differentiation of the concepts of difficulty and ability. *Child Development, 54,* 951-959. DOI: 10.2307/1129899

Nicholls, J. G., & Miller, A. T. (1984a). Development and its discontents: The differentiation of the concept of ability. In J. G. Nicholls (Ed.), *Advances in motivation and achievement: The development of achievement motivation* (pp. 185–218). Greenwich, CT: JAI Press.

Nicholls, J. G., & Miller, A. T. (1984b). Reasoning about the ability of self and others: A developmental study. *Child Development, 55,* 1990-1999. DOI: 10.2307/1129774

Parsons, J. E., & Ruble, D. N. (1977). The development of achievement-related expectancies. *Child Development, 48,* 1075-1079. DOI: 10.2307/1128364

Preacher, K. J., Curran, P. J., & Bauer, D. J. (2006). Computational tools for probing interactions in multiple linear regression, multilevel modeling, and latent curve analysis. *Journal of Educational and Behavioral Statistics, 31,* 437-448. https://doi.org/10.3102/10769986031004437

Rholes, W. S., Blackwell, J., Jordan, C., & Walters, C. (1980). A developmental study of learned helplessness. *Developmental Psychology, 16,* 616-624. http://dx.doi.org/10.1037/0012-1649.16.6.616

Rholes, W. S., & Ruble, D. N. (1984). Children's understanding of dispositional characteristics of others. *Child Development*, *55*, 550-560. DOI: 10.2307/1129966

Ruble, D. N., Boggiano, A. K., Feldman, N. S., & Loebl, J. H. (1980). Developmental analysis of the role of social comparison in self-evaluation. *Developmental Psychology*, *16*, 105-115. http://dx.doi.org/10.1037/0012-1649.16.2.105

Rosenholtz, S. J., & Rosenholtz, S. H. (1981). Classroom organization and the perception of ability. *Sociology of Education*, *54*, 132-140. DOI: 10.2307/2112357

Schuster, B., Ruble, D. N., & Weinert, F. E. (1998). Causal inferences and the positivity bias in children: The role of the covariation principle. *Child Development*, *69*, 1577-1596. https://doi.org/10.1111/j.1467-8624.1998.tb06178.x

Senko, C., Hulleman, C. S., & Harackiewicz, J. M. (2011). Achievement goal theory at the crossroads: Old controversies, current challenges, and new directions. *Educational Psychologist*, *46*, 26-47. https://doi.org/10.1080/00461520.2011.538646

Stipek, D. J. (1981). Children's perceptions of their own and their classmates' ability. *Journal of Educational Psychology*, *73*, 404-410. http://dx.doi.org/10.1037/0022-0663.73.3.404

Smiley, P. A., & Dweck, C. S. (1994). Individual differences in achievement goals among young children. *Child Development*, *65*, 1723-1743. https://doi.org/10.1111/j.1467-8624.1994.tb00845.x

Stipek, D. J., & Daniels, D. H. (1988). Declining perceptions of competence: A consequence of changes in the child or in the educational environment? *Journal of Educational Psychology*, *80*, 352-356. http://dx.doi.org/10.1037/0022-0663.80.3.352

Stipek, D. J., & Hoffman, J. M. (1980). Development of children's performance-related judgments. *Child Development*, *51*, 912-914. DOI: 10.2307/1129485

Thomaes, S., Brummelman, E., & Sedikides, C. (2017). Why most children think well of themselves. *Child Development*, *88*, 1873-1884. https://doi.org/10.1111/cdev.12937

Weiner, B. (1985). An attributional theory of achievement motivation and emotion. *Psychological Review*, *92*, 548–573. http://dx.doi.org/10.1037/0033-295X.92.4.548

Wellman, H. M., & Gelman, S. A. (1992). Cognitive development: Foundational theories of core domains. *Annual Review of Psychology*, *43*, 337-375. DOI: 10.1146/annurev.ps.43.020192.002005

Wigfield, A, Eccles, J. S., Fredricks, J., Simpkins, Roeser, R., & Schiefele, U. (2015). Development of achievement motivation and engagement. In R. Lerner (Series ed.) and M. Lamb & C. Garcia Coll (Vol. Eds.), *Handbook of child psychology and developmental science* (7th ed., pp. 657–700). New York, NY: Wiley.

Yang, F., & Frye, D. (2016). Early understanding of ability. *Cognitive Development*, *38*, 49-62. https://doi.org/10.1016/j.cogdev.2016.01.003

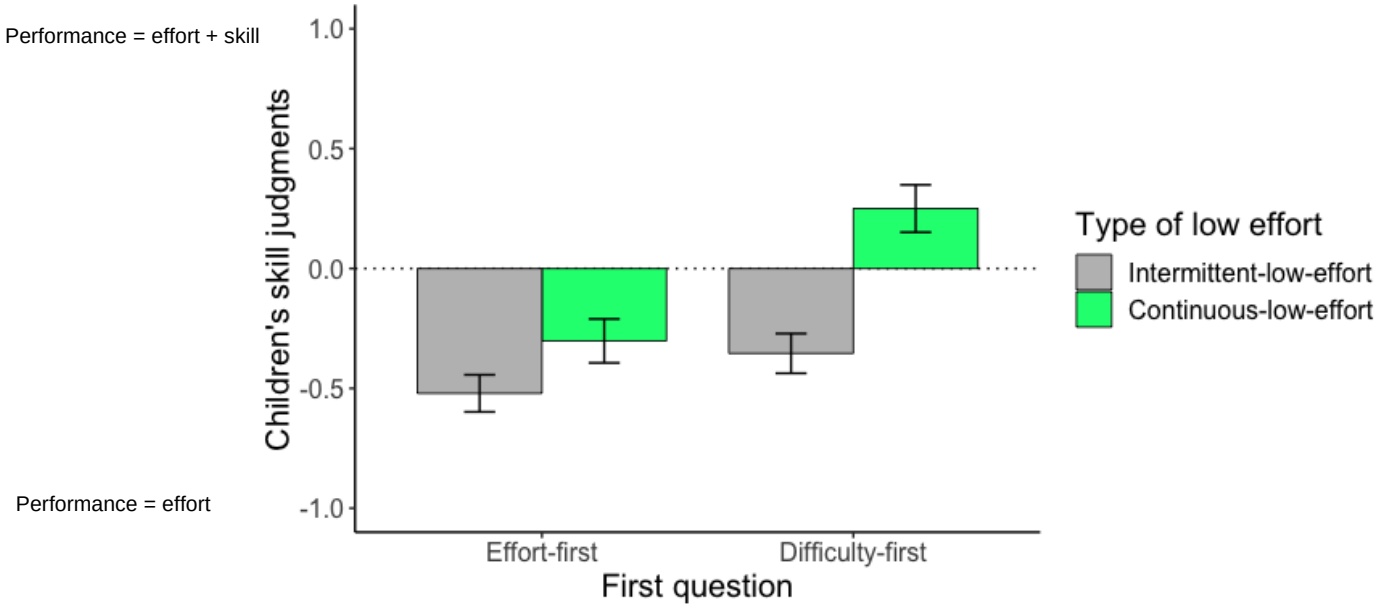

Figure 1. Observed means of children's skill judgments (across the evaluation and prediction questions) in Experiment 1, plotted by the content of the first question (effort-first vs. difficulty-first) and the type of low effort (intermittent- vs. continuous-low-effort). Error bars represent ± 1 *SE*.

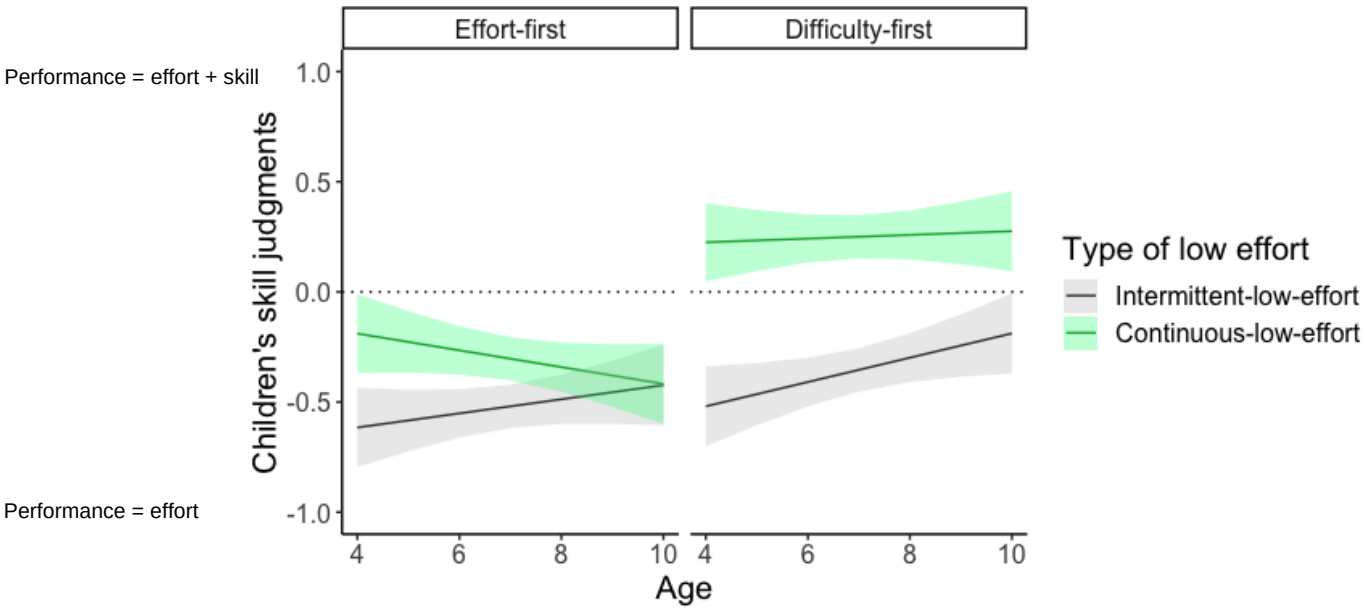

Figure 2. Marginal means of children's skill judgments (across the evaluation and prediction questions) in Experiment 1, plotted by age, the content of the first question (effort-first vs. difficulty-first), and the type of low effort (intermittent- vs. continuous-low-effort). Confidence bands represent ± 1 *SE*.

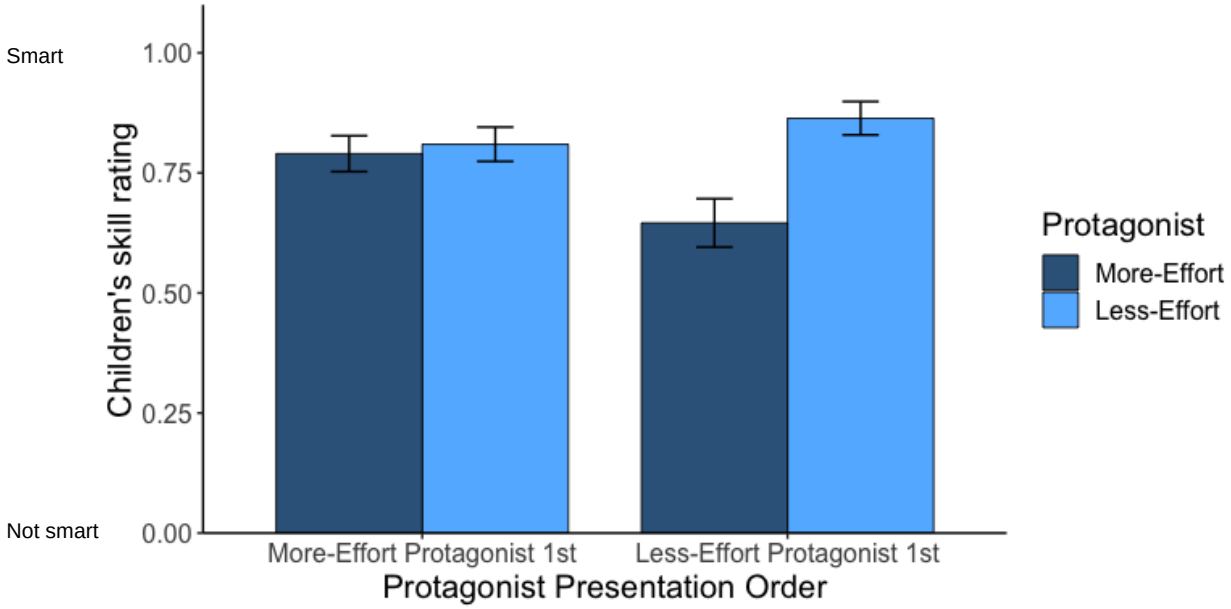

Figure 3. Observed means of children's skill ratings (across the evaluation and prediction questions) in Experiment 2, plotted by protagonist (more-effort vs less-effort character) and the presentation order of the two characters (more-effort protagonist first vs. less-effort protagonist first). Error bars represent ± 1 *SE*.

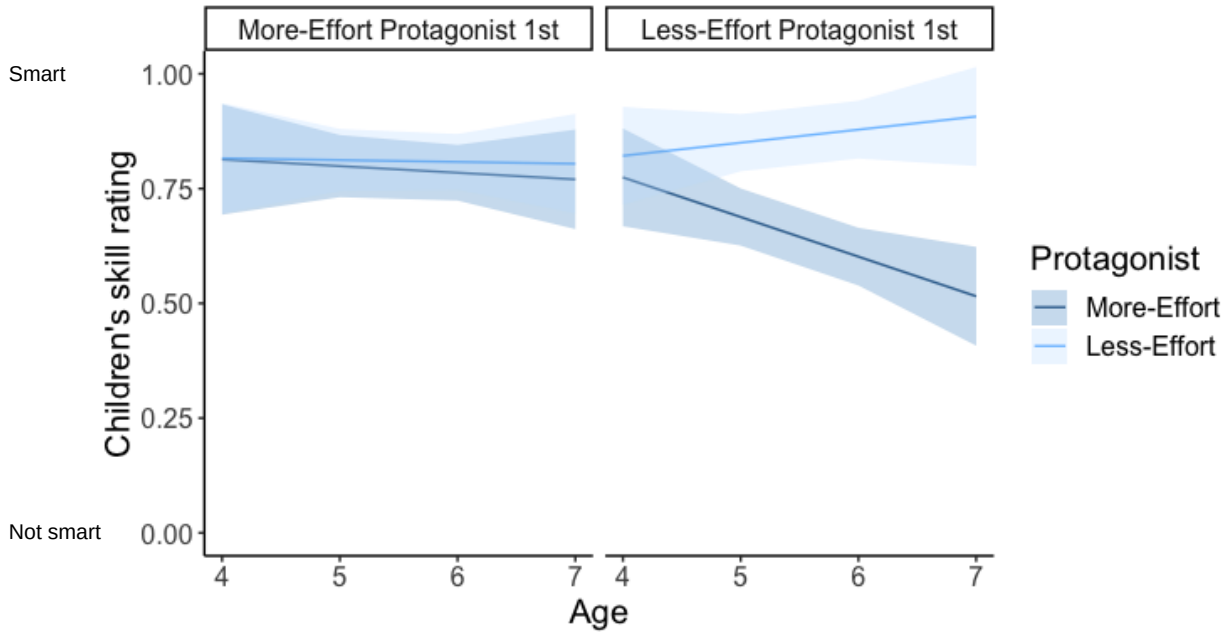

Figure 4. Marginal means of children's skill ratings (across the evaluation and prediction questions) in Experiment 2, plotted by age, protagonist (more-effort vs less-effort protagonist) and the presentation order of the two characters (more-effort protagonist first vs. less-effort protagonist first). Confidence bands represent ± 1 *SE*.