

The title for this Special Section is **Origins of Children's Self-Views**, edited by Eddie Brummelman and Sander Thomaes

## When Peer Performance Matters: Effects of Expertise and Traits on Children's Self-Evaluations After Social Comparison

Candace Lapan and Janet J. Boseovski  
*The University of North Carolina at Greensboro*

The present research examined the influence of peer characteristics on children's reactions to upward social comparisons. In Experiment 1, one hundred twenty-six 5-, 8-, and 10-year-olds were told that they were outperformed by an expert or novice peer. Older children reported higher self-evaluations after comparisons with an expert rather than a novice, whereas 5-year-olds reported high self-evaluations broadly. In Experiment 2, ninety-eight 5- to 6-year-olds and 9- to 10-year-olds were told that the peer possessed a positive or negative trait that was task relevant (i.e., intelligence) or task irrelevant (i.e., athleticism). Older children reported higher self-evaluations after hearing about positive rather than negative traits, irrespective of relevance. Younger children reported high self-evaluations indiscriminately. Results inform the understanding of social comparison development in childhood.

Self-evaluations of performance influence children's academic motivation and achievement (Schunk & Pajares, 2002). Beginning in middle childhood, social comparison with peers becomes a potent source of input for self-evaluations (Ruble, Boggiano, Feldman, & Loebel, 1980). Moreover, information about peers moderates the effects of social comparison on self-evaluations (e.g., gender; Rhodes & Brickman, 2008). The current studies examined whether children's responses to social comparison depend on the expertise (i.e., knowledge) and ability-related traits (e.g., smartness) of a peer. These factors were chosen because they are salient in childhood and they provide a direct basis for making judgments of competence (e.g., children should feel better about themselves when outperformed by an expert rather than a novice).

By 5 years of age, children understand relative performance feedback (Morris & Nemcek, 1982) and show interest in peers' academic work (Frey & Ruble, 1985), but viewing peers' work has little influence on

self-evaluations until 7 years of age (Ruble et al., 1980). In middle childhood, worldviews become more realistic and self-evaluations are altered by negative feedback. Indeed, 7-year-olds' self-evaluations are lower when upward comparative feedback, rather than an objective performance standard, is provided (Ruble et al., 1980). In late childhood, interest in peers' work remains high, but upward comparisons impair motivation (Butler, 1989), affect (Ruble et al., 1980), and self-evaluations (Pomerantz, Ruble, Frey, & Greulich, 1995).

Little developmental research has investigated the effects of peer characteristics on children's self-evaluations in a social comparison setting. Generally, preschoolers infer that individual differences in behavior (Rhodes & Gelman, 2008) and ability (Cimpian, 2010) are tied to social categories (e.g., gender). Such characteristics may be used to interpret performance differences. Indeed, when preschoolers experience failure relative to an opposite gender peer, they report low self-evaluations (Cimpian, 2010; Rhodes & Brickman, 2008). Older children are sensitive to subtler peer characteristics. For example, 9- to 11-year-olds' academic self-concept is impacted by comparisons with reciprocated

---

This research was conducted in partial fulfillment of the doctoral degree of C. Lapan. We are grateful to the children, parents, and teachers who supported this project. The authors also thank Valeriya Tsygankova, Carmin Ball, and Kimberly Marble for their assistance with both studies.

Correspondence concerning this article should be addressed to Janet J. Boseovski, Department of Psychology, UNCG, PO Box 27402-6170, Greensboro, NC. Electronic mail may be sent to [jjboseov@uncg.edu](mailto:jjboseov@uncg.edu).

© 2017 The Authors

Child Development © 2017 Society for Research in Child Development, Inc. All rights reserved. 0009-3920/2017/8806-0009

DOI: 10.1111/cdev.12941

friends, but not with nonreciprocated friends (Altermatt & Pomerantz, 2003). Also, social comparisons have greater influence on 9- and 10-year-olds' self-perceptions when peers are described as competitors (Butler, 1989).

Children's emerging ability conceptions are thought to account, in part, for age-related sensitivity to social comparison (Butler, 1992). Younger children's lack of receptivity to feedback may stem from the belief that failure is due to temporary factors (e.g., effort; Nicholls, 1978) rather than stable differences in ability (e.g., smartness). In later childhood, greater sensitivity to comparative feedback is related to older children's increased tendency to interpret relative failure as a result of differences in fixed traits (e.g., intelligence; Butler, 1989). Indeed, first and third graders report that people can change their abilities if they want to, whereas older children and adults tend to report that abilities cannot be changed (Gelman, Heyman, & Legare, 2007, Experiment 2). There are also direct links between fixed ability conceptions and social comparison behaviors. For example, task descriptions that emphasize fixed ability (e.g., indicating that the task assesses ability rather than performance) increase sixth graders' interest in comparative feedback (Butler, 1992). In middle childhood, children who frequently endorse beliefs of ability as constant seek out more comparative feedback than children who endorse these beliefs less often (Ruble & Flett, 1988). Further, these developments in children's ability conceptions and interpretations of feedback are linked to changing parental and educational socialization practices between early and middle childhood (for review, see Cimpian, 2017).

Given the links between ability conceptions and social comparison, the provision of peer characteristics that are related to competence should alter children's interpretations of relative failure and, thereby, self-evaluations. The current studies examined two relevant aspects of ability: expertise and ability-related traits. Expertise refers to possession of a specialized body of knowledge (Landrum, Mills, & Johnston, 2013). Conversely, ability-related traits are thought of as relatively fixed and central characteristics of individuals (Lockhart, Goddu, & Keil, 2016).

There are two main similarities in the developmental trajectories of expertise and ability-related trait conceptualizations. First, children use expertise and ability-related traits to infer individuals' knowledge early in life, and these characteristics are seen as increasingly interrelated with age. Young children recognize that experts have superior

knowledge in their domain of expertise compared with laypeople (Landrum et al., 2013). For example, given conflicting information, even preschoolers endorse information provided by a person described as an expert rather than a layperson (e.g., Koenig & Jaswal, 2011). Concerning ability-related traits, kindergarteners cite amount of knowledge when defining "smart" and this tendency increases with age (Kurtz-Costes, McCall, Kinlaw, Wiesen, & Joyner, 2005). Also, even 4-year-olds believe that an actor who provided appropriate object names is smarter than one who lifted heavy objects (Hermes, Behne, & Rakoczy, 2015). By 7 years of age, children use ability-related traits to infer specific knowledge (e.g., smart vs. nice reflect different knowledge bases; Danovitch & Keil, 2007).

Second, between early and middle childhood, both expertise and ability-related traits are seen as increasingly relevant to performance outcomes. Although preschoolers use expertise to make decisions about the correctness of information, it is not until middle childhood that expertise guides children's selections of a task instructor (Boseovski, Hughes, & Miller, 2016). This suggests a developing awareness that experts should perform well on tasks that are relevant to the domain of expertise. Similarly, with age, children perceive ability-related traits as linked to achievement. Five- to 8-year-olds spontaneously explain academic outcomes by actions (e.g., studies a lot), whereas 10-year-olds use both action and trait explanations (e.g., "He's smart"; Benenson & Dweck, 1986). The emerging salience of these cues suggests that peer expertise and ability-related traits should be relevant to children's interpretation of comparative feedback.

Despite the developmental salience of expertise and ability-related traits, it is possible that the use of these cues will be limited in self-evaluative situations. Although there is no direct evidence for this assertion, research suggests that children have difficulty reasoning about these characteristics in contexts that involve the self as compared with others. For instance, 4- and 5-year-olds can select the most relevant expert to consult in a particular domain, but they fail to do so when given the option to rely on their own knowledge (Aguiar, Stoess, & Taylor, 2012, Experiments 1 and 2). In contrast, 6-year-olds recognize limitations in their own knowledge. Trait reasoning is also affected by self-evaluative contexts: children reason more aptly about covariation information when making attributions about others than the self, and this effect decreases between 5 and 9 years of age (Schuster, Ruble, & Weinert, 1998). Thus, the self-evaluative nature of social

comparison may also limit children's reasoning about these cues.

Developmental change in the processing of positive information (i.e., positivity bias) also affects expertise and ability-related trait conceptions (Lan-drum et al., 2013, Experiments 2 and 3). For example, 6- to 7-year-olds are less willing to accept negative information about a novel animal (e.g., "dangerous") as compared with positive information (e.g., "friendly"), even when the information is given by a zookeeper (Boseovski & Thurman, 2014). Thus, positive information can override children's reliance on expertise cues. Similarly, spontaneous ability-related trait explanations emerge earlier for success than failure (Benenson & Dweck, 1986). Of relevance here, a desire to view themselves positively may also override children's consideration of peer expertise and ability-related traits in a social comparison setting.

Although expertise and ability-related trait reasoning follow similar developmental trajectories, the extent to which children essentialize these characteristics diverges with age. Essentialism reflects the belief that characteristics are biologically based, present at birth, stable over time, and immune to change (Gelman et al., 2007). Essentialist beliefs about ability-related traits are evident in some children by 5 years of age (Heyman & Gelman, 2000), become more prevalent with age, and are more common than essentialist beliefs about personality traits. For example, 7- to 14-year-olds require less evidence to attribute ability-related traits than personality traits (e.g., "nice"; Heyman & Giles, 2004). Between 5 and 10 years of age, children increasingly expect ability-related traits, as compared with personality traits, to be linked to nature (e.g., believe that adopted children will share traits with their birth parents, not adopted parents; Heyman & Gelman, 2000). By late childhood, children tend to view ability-related traits as unchangeable (Gelman et al., 2007), although there are substantial individual differences in these beliefs.

In contrast, expertise has not been conceptualized in a framework of essentialism, as knowledge has the potential to change (Lockhart et al., 2016). Thus, methodological assessments of expertise differ from those that assess perceptions of ability-related traits. In many studies, expertise is denoted by referring to individuals' experience (e.g., "has taken many classes on sewing"; Boseovski et al., 2016). Assessments of expertise reasoning evaluate children's understanding of the amount of knowledge an individual has (e.g., "How much would *he* know when he is 30 years old?"; Lockhart et al.,

2016), whereas assessments of ability-related traits evaluate children's reasoning about the essence of an individual (i.e., "Could he change to become *a smart person* if he wanted to?"; Gelman et al., 2007). Findings indicate that children conceptualize expertise as largely malleable throughout early and middle childhood. For example, 5- to 12-year-olds expect that an individual can change how much they know and expect a faster rate of knowledge growth than adults (Lockhart et al., 2016, Experiment 1).

The current studies examined the effects of peer expertise (Experiment 1) and ability-related traits (e.g., smartness; Experiment 2) on 5- to 10-year-olds' self-evaluations in a comparison setting. Concerning general predictions for Experiment 1, we expected that younger children would report high self-evaluations regardless of peer expertise, and that with age, children would report lower self-evaluations after being outperformed by a novice rather than an expert. In Experiment 2, we predicted that younger children may be more sensitive to traits than expertise, given the salience of trait categories early in life. Older children were further expected to differentiate their self-evaluations based on trait relevance (e.g., reporting lower self-evaluations when outperformed by an unintelligent peer, but not a nonathletic peer), as described in the following section. This examination provides developmental insight about the use of these cues in self-relevant contexts across an age range in which negative feedback is seen as increasingly relevant. Also, the attenuation of the positivity bias was expected to be reflected in lower self-evaluations with age.

### Experiment 1

To our knowledge, no study has examined children's use of expertise information in a social comparison setting. We assessed 5-, 8-, and 10-year-olds' self-evaluations after they were told that they were outperformed by an expert peer as compared with a novice peer on a novel task and a familiar task. Drawing was selected as a commonplace activity, whereas the Tower of Hanoi (ToH) task (Welsh, 1991) was expected to be unfamiliar to children.

Comparative feedback should have greater effects on self-evaluations with age given developmental shifts in ability conceptions (Butler, 1989) and socialization practices (Cimpian, 2017). We expected 5-year-olds' self-evaluations to be high regardless of peer expertise, given their general

optimism (Boseovski, 2010) and limited expertise reasoning in self-relevant contexts (Aguiar et al., 2012). It was unclear how 8-year-olds would respond to peer expertise, as they are sensitive to comparative feedback (Ruble et al., 1980), but sometimes display a positivity bias (Boseovski & Thurman, 2014). We expected 10-year-olds' self-evaluations to be lower after being outperformed by a novice as opposed to an expert, although the effects may be weak given children's perceptions of expertise as malleable. Because comparative feedback has stronger effects on self-evaluations in domains of personal significance (Bers & Rodin, 1984), we expected these effects to be more robust in a familiar domain, as drawing may be perceived as more personally relevant than an unfamiliar task. These predictions applied to children's affect and performance evaluations. Ability evaluations were expected to be unaffected, as multiple comparisons are often necessary to influence these evaluations (Ruble et al., 1980).

### Method

#### Participants

Forty-one 5-year-olds ( $M = 66.2$  months,  $SD = 3.7$ , 23 males), forty-three 8-year-olds ( $M = 102.2$  months,  $SD = 3.2$ , 19 males), and forty-two 10-year-olds ( $M = 126.7$  months,  $SD = 3.6$ , 24 males) were tested between March 2015 and March 2016. Concerning demographic data, 74.6% of children were Caucasian, 15.9% African American, 7.9% mixed race, 0.8% Asian; 0.8% of families did not disclose this information. Family income ranged from < \$20,000 to over \$90,000.

#### Materials

In the drawing task, an  $8.5 \times 11$  in. line drawing of a cat was used. Task feedback was given on a 13 in. Macbook Pro (Apple Inc., One Infinite Loop, Cupertino, California) and displayed via PowerPoint slides. A cardboard face with an adjustable mouth was used to assess children's affective states (Ruble et al., 1980).

#### Design and Procedure

A mixed design was used to assess the effects of age (between-subjects: 5-, 8-, and 10-year-olds), comparison target expertise (between-subjects: expert vs. novice), and domain (within-subjects: novel vs. familiar) on children's self-evaluations. As

part of a larger study, children were tested in the university laboratory or their day care facility by a female experimenter in one session. Task order was counterbalanced between participants.

*Novel domain.* Participants were presented with two stacking toys and the experimenter described the task as the "flerping game." Children were told that the goal was to get all of their discs to the last post in increasing size (for procedure, see Welsh, 1991). Children completed three practice trials and received corrective feedback. Then, they were told "You're going to play the flerping game for 3 min. Try to get as many as you can and try not to make mistakes." Time was not actually limited (i.e., they played until they completed five trials), but was introduced as a supposed factor to increase performance ambiguity. Then, the experimenter pretended to enter information in the computer saying,

I'll put in how many moves you made and how long it took you to finish. Then, the computer will tell us how you did on the flerping game. You'll get a score in stars. The more stars you get, the better you did.

The experimenter showed children a display of nine stars and said "You got nine stars on the flerping game."

Next, children were told that the computer had other children's scores as well. Based on their pre-assigned condition, children received a description of either an expert or novice peer matched to the participants' gender. For example, in the expert condition, children were told:

Casey is a boy your age . . . He knows a lot about flerping. After school, he often takes special lessons on how to flerp. He knows how to flerp with lots of discs and posts. He flerps in front of lots of people. He has won competitions for his flerping.

The experimenter then showed children a visual display and stated, "Casey got 11 stars on the flerping game." The scores of 9 and 11 were selected to avoid the suggestion of a normative maximum (e.g., scores of 5 and 7 may suggest that the maximum was 10; see Appendix S1 for full descriptions).

We assessed children's beliefs about whether the peer's level of expertise influenced their performance. Peer performance questions were based on condition (expert: "Did Casey get 11 stars because he

knew a lot about flerping?"; novice: "Did Casey get 11 stars because he knew a little about flerping?"). "Yes" responses were coded as 1 and "no" responses were coded as 0. Next, children answered three self-evaluation questions (Ruble, Eisenberg, & Higgins, 1994). First, they were asked, "How do you feel about how you did on the flerping game?" and manipulated the mouth on the face to indicate their affective state (marks allowed for scoring; 1 = *very sad*, 17 = *very happy*). Second, they were asked to evaluate their performance, "How do well do you think you did on the flerping game?" and ability, "How good do you think you are at doing games like the flerping game?" on a 9-point Likert scale accompanied by a display of circles of increasing size (1 = *not good at all*, 9 = *very good*). Children provided open-ended explanations for their self-evaluations (Appendix S2), which were coded by a one rater, and another rater coded 20%. High interrater reliability was attained all variables ( $\kappa_s = .84-.91$ ).

*Familiar domain.* For the drawing task, children were shown a picture of a cat and told

We're going to play a drawing game. You have to try and draw this picture on your own. You will get 3 min to draw it. Try and make your drawing as close as possible to this one and try not to make mistakes.

Again, time to complete the task was not really limited. Next, the experimenter took a digital picture of the child's drawing and said, ". . . the computer will grade it and tell us how close your picture was to this one." Again, children heard about either expert or novice peers. For instance, in the novice condition, children were told

Lee is a boy your age . . . Lee knows a little about drawing. He has taken the art class at school that all kids take, but no other special art classes outside of school. He knows how to draw a few things. He only shows his art to a few people. He's never put his art in a competition to try and win a prize.

As with the novel task, children were told that they received 9 stars and the peer received 11 stars.

### Results

One 10-year-old female was excluded from data analyses because she was familiar with the ToH task. One 5-year-old male was excluded because he failed the practice trials of the ToH.

### Peer Performance Question

Children of all ages reliably indicated that the expert peer got 11 stars because he or she "knew a lot" about the domain (drawing: 100%); *t*-tests against chance indicated that this was true for flerping (96%),  $p < .05$ . In the novice condition, 8- and 10-year olds reliably rejected the idea that the novice peer got 11 stars because they "knew a little" about the domain (drawing: 82.9%; flerping 82.9%,  $ps < .05$ ), but 5-year-olds' responses did not differ significantly from chance in either domain: (drawing: 60%; flerping, 52.6%),  $ps > .10$ .

### Self-Evaluations

Descriptive data are presented in Table 1. Data were analyzed using a series of 3 (age: 5-year-olds vs. 8-year-olds vs. 10-year-olds; between-subjects)  $\times$  2 (target expertise: expert vs. novice; between-subjects)  $\times$  2 (domain: familiar vs. novel; within-subjects) mixed analyses of variance. There were no significant effects or interactions involving gender or task order on any dependent measures; thus, these variables were excluded from the final models.

*Affect.* Five-year-olds felt better about their performance than older children,  $F(2, 120) = 7.27$ ,  $p = .001$ ,  $\eta_p^2 = .11$ . There was a significant interaction between age and domain,  $F(2, 120) = 3.62$ ,  $p = .03$ ,  $\eta_p^2 = .06$ . Ten-year-olds' affective ratings varied significantly by domain,  $F(1, 40) = 5.71$ ,  $p = .02$ ,  $\eta_p^2 = .13$ , such that affective ratings were more positive in the novel domain ( $M = 14.44$ ,  $SD = 2.40$ ) than the familiar domain ( $M = 13.51$ ,  $SD = 2.75$ ),  $t(41) = -2.33$ ,  $p = .03$ . Conversely, 5- and 8-year-olds' affective ratings did not differ by domain,  $ps > .05$ , and were relatively positive (Table 1). No other main effects or interactions were significant,  $ps > .05$ .

*Performance.* There were significant main effects of age,  $F(2, 120) = 5.94$ ,  $p = .003$ ,  $\eta_p^2 = .09$ , and expertise,  $F(1, 120) = 5.04$ ,  $p = .03$ ,  $\eta_p^2 = .04$ . These effects were qualified by a significant interaction between age, peer expertise, and domain,  $F(2, 120) = 2.93$ ,  $p = .05$ ,  $\eta_p^2 = .05$ ; see Figure 1a. Five-year-olds' performance evaluations were relatively high and did not differ by peer expertise or domain,  $ps > .05$ . Eight-year-olds reported lower evaluations when they were outperformed by a novice rather than an expert, irrespective of domain,  $F(1, 41) = 7.55$ ,  $p = .01$ ,  $\eta_p^2 = .16$ . Ten-year-olds' evaluations varied by domain and expertise,  $F(1, 40) = 5.83$ ,  $p = .02$ ,  $\eta_p^2 = .13$ ; 10-year-olds reported lower evaluations after being outperformed by a

Table 1  
 Mean Self-Evaluations in Experiment 1 (by Age, Domain, and Peer Expertise) and Experiment 2 (by Age and Trait Valence)

Experiment 1																		
Familiar task																		
Age in years	Expert									Novice								
	Affect			Performance			Ability			Affect			Performance			Ability		
	M	SE	N	M	SE	N	M	SE	N	M	SE	N	M	SE	N	M	SE	N
5	15.60	0.60	20	8.00	0.47	20	8.80	0.11	20	16.23	0.40	21	8.23	0.26	21	8.47	0.25	21
8	14.19	0.53	21	7.52	0.25	21	7.47	0.38	21	13.09	0.99	22	6.73	0.31	22	8.00	0.23	22
10	14.00	0.79	21	7.42	0.34	21	7.25	0.43	21	13.14	0.62	21	6.52	0.25	21	7.05	0.36	21
Novel task																		
Age in years	Expert									Novice								
	Affect			Performance			Ability			Affect			Performance			Ability		
	M	SE	N	M	SE	N	M	SE	N	M	SE	N	M	SE	N	M	SE	N
5	15.30	1.33	20	8.25	0.45	20	8.60	0.21	20	14.80	0.64	21	7.57	0.48	21	7.90	0.37	21
8	13.52	0.85	21	7.76	0.24	21	7.95	0.25	21	11.77	0.99	22	6.68	0.31	22	7.36	0.36	22
10	14.29	0.54	21	7.47	0.25	21	7.38	0.39	21	14.71	0.54	21	7.52	0.24	21	7.81	0.27	21
Experiment 2																		
Age in years	Positive trait									Negative trait								
	Affect			Performance			Ability			Affect			Performance			Ability		
	M	SE	N	M	SE	N	M	SE	N	M	SE	N	M	SE	N	M	SE	N
5-6	14.76	0.85	25	7.76	0.45	25	8.44	0.32	25	15.32	0.57	25	7.80	0.45	25	8.24	0.22	25
9-10	14.79	0.44	24	7.42	0.20	24	7.46	0.26	24	10.88	0.70	24	6.04	0.19	24	6.92	0.24	24

novice rather than an expert in the familiar domain,  $t(40) = -2.06, p = .05$ , but not the novel domain,  $t(40) = 0.14, p = .89$ . Eight- and 10-year-olds' evaluations after expert comparisons were high; indeed, they were not significantly different from 5-year-olds' evaluations in either domain,  $ps > .10$ .

**Ability.** Ability ratings decreased with age,  $F(2, 120) = 11.49, p < .001, \eta_p^2 = .16$ , but were quite high (Table 1). No other main effects or interactions were significant,  $ps > .05$ .

*Explanations*

Descriptive data and analyses of explanations are provided in Appendices S3 and S4, respectively.

*Discussion*

These findings are the first to reveal age-related change in children's use of expertise information in

a social comparison setting. Five-year-olds' self-evaluations were positive despite relative failure. In contrast, 8- and 10-year-olds showed lower performance evaluations in response to feedback, but this was limited to the familiar domain for the latter age group. Notably, effects of peer expertise were seen only for performance evaluations and had no significant effect on children's reported feelings about their performance or ability evaluations.

Five-year-olds' positive self-evaluations are consistent with previous research (Ruble et al., 1994). Despite findings that children of this age use expertise information readily in many contexts (e.g., Koenig & Jaswal, 2011), our findings reveal that it had limited utility for young children in this specific context. These results are also consistent with research documenting a positivity bias in this age group (Boseovski, 2010). The self-relevance of the comparison context may have affected young children's recognition that a novice has little task

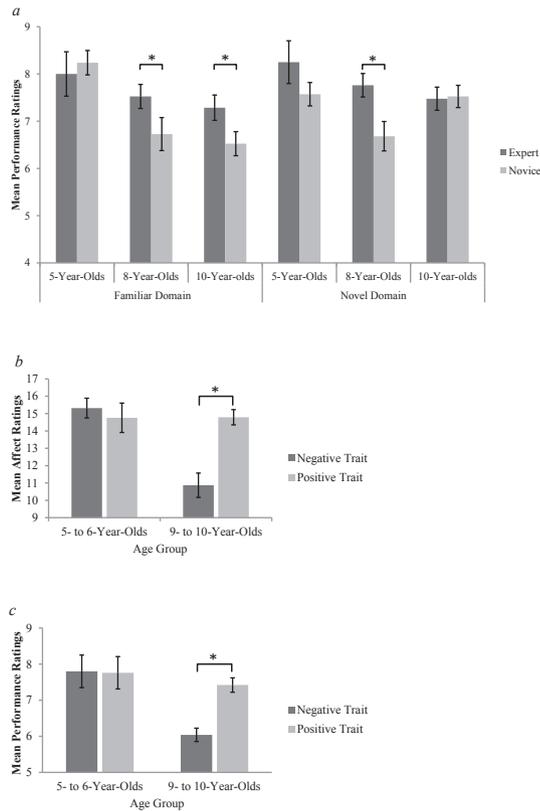


Figure 1. Mean self-evaluations in Experiment 1 (by age, domain, and peer expertise) and Experiment 2 (by age and trait valence). Error bars represent standard errors. (a) Mean performance evaluations in Experiment 1 by age, domain, and peer expertise. (b) Mean affect ratings in Experiment 2 by age and trait valence. (c) Mean performance evaluations in Experiment 2 by age and trait valence.

\* $p < .05$ .

knowledge and that relative failure in this case indicates particularly poor performance. As discussed earlier, preschoolers have difficulty reasoning about knowledge when judgments involve the self (Aguiar et al., 2012), which is likely related to inferences about relative failure in this study. On the peer performance questions, 5-year-olds understood the relation between high expertise and successful performance, but were unsure about the relation between low expertise and performance. Thus, failure relative to a novice was unlikely to result in poor performance evaluations.

Unlike younger children, older children's performance evaluations varied appropriately by peer expertise. As predicted, they viewed their performance more poorly when they were outperformed by a novice rather than an expert, revealing an understanding that failure relative to an expert was not a meaningful indicator of their performance. It was surprising, however, that 8- and 10-year-olds'

feelings about their performance did not vary by expertise. There is mixed evidence for the consistency of performance evaluations and affective responses to social comparison (Ruble et al., 1994), and this dissociation warrants further research. Expertise information likely did not affect emotional reactions to relative failure because it is seen as relatively malleable and not highly evaluative (e.g., being a novice is not necessarily a negative quality, in contrast to incompetence; Koenig & Jaswal, 2011). Highly evaluative peer attributes that are more essentialized (e.g., ability-related traits) may elicit stronger emotional reactions.

As expected, children's reactions to social comparison varied by domain. Surprisingly, 8-year-olds' differentiated their performance evaluations by peer expertise in both domains, whereas 10-year-olds only did so in the familiar domain. Eight-year-olds may not yet clearly distinguish tasks based on personal significance, which can lead to distinct responses to feedback. Conversely, 10-year-olds may have devalued the novel task and thus, perceived comparative feedback as less important (Bers & Rodin, 1984). By late childhood, children may value their personal creation (i.e., their drawing) more than their fairly abstract performance on an unfamiliar task. Indeed, adults and children place particularly high value on original creations (Frazier & Gelman, 2009). Although older children showed sensitivity to comparative feedback in their quantitative self-evaluations in the familiar domain, both these children and younger children explained their performance in positive terms (e.g., "It looks really good").

In Experiment 2, comparison peers were labeled in trait terms. Given the mixed findings from Experiment 1 in the novel domain, Experiment 2 included only a novel domain to examine these effects further. Also, we assessed these effects in 5- to 6-year-olds and 9- to 10-year-olds, as these are pivotal ages for trait reasoning development (Heyman & Gelman, 2000).

## Experiment 2

To our knowledge, this is the first experiment to explore children's use of peer trait labels in social comparison. Given the inductive nature of trait labels (Heyman & Gelman, 2000), the early association of "smartness" with success (Kurtz-Costes et al., 2005), and age-related increases in views of ability as fixed, we expected trait information to be more influential than expertise in this context. Also, older children may need such information for

relative failure to be deemed relevant in a novel domain. Five- to 6-year-olds and 9- to 10-year-olds were told that they were outperformed on a novel task by a peer with a positive or negative trait that was task-relevant (i.e., intelligence) or task-irrelevant (i.e., athleticism). These traits were chosen because they are understood by young children (Lockhart, Keil, & Aw, 2013). We included relevant and irrelevant traits to assess whether young children apply traits inaptly. Research indicates that 5-year-olds, but not 10-year-olds, overgeneralized the importance of being smart to irrelevant tasks (i.e., jumping hurdles; Kurtz-Costes et al., 2005). Piloting indicated that athleticism was seen as relevant to ToH performance given motor demands of the task (i.e., moving the discs). Thus, we used Raven's Coloured Progressive Matrices (RCPM), which does not involve motor responses.

We expected older children to report lower affect and performance evaluations after being outperformed by a peer with a relevant negative trait, but to be unaffected by peers labeled with irrelevant traits (i.e., athleticism). For 5- to 6-year-olds, one possibility is that traits would have little impact on self-evaluations, as in Experiment 1, due in part to a positivity bias and limited trait reasoning in self-relevant contexts. It is also possible that these children would be more sensitive to traits given their strong salience in early childhood (Kurtz-Costes et al., 2005). For all participants, ability evaluations were expected to be high. Finally, we assessed whether children attributed academic outcomes to ability-related traits (i.e., entity reasoning) or malleable factors (i.e., incremental reasoning), which is one aspect of trait essentialism (Gelman et al., 2007).

### Method

#### Participants

Fifty 5- to 6-year-olds ( $M = 72.4$  months,  $SD = 7.4$ , 28 males) and forty-eight 9- to 10-year-olds ( $M = 119.9$  months,  $SD = 6.9$ , 25 males) were recruited in the same manner and time frame as Experiment 1. For demographics, 62.2% of participants were Caucasian, 19.4% African American, 11.2% mixed races, and 2% Hispanic. Also, 5.1% of families chose not to report race or ethnicity. Household incomes ranged from < \$20,000 to over \$90,000.

#### Materials

Children completed the RCPM booklet (Raven, Court, & Raven, 1995).

#### Design and Procedure

A between-subjects design was used to assess the effects of age (5- to 6-year-olds vs. 9- to 10-year-olds), trait valence (positive vs. negative), and trait relevance (relevant vs. irrelevant) on children's self-evaluations. Children completed the Entity/Incremental task, followed by the social comparison task.

*Entity/incremental task.* Children were asked two questions about academic success (self and other) and two questions about academic failure (self and other; adapted from Benenson & Dweck, 1986). For example, children were asked "Think of kids in your class who get a lot wrong on their schoolwork. Why do they get a lot wrong?" Process responses were coded as 0, and Ability responses were coded as 1 (Appendix S5). The number of Ability responses was summed across the four questions; thus, the possible range of scores was 0–4.

*Social comparison task.* Children were shown the RCPM and given standard instructions (Raven et al., 1995), but were told that they had 5 min to complete as many items as possible. As in Experiment 1, children were actually given as much time as needed to complete the task. Next, the experimenter said, "I'll put your answers and how long it took you to finish into the computer. Then, the computer will tell us how you did on the matrix game." Feedback procedures were identical to Experiment 1 except for the peer characteristics provided. Children either heard about a peer with a positive trait or a peer with a negative trait. Further, the trait was either task relevant (i.e., smart/not smart) or task irrelevant (i.e., athletic/not athletic; see Appendix S6). For example, children in the positive relevant condition heard "Casey is a smart boy. He knows lots of things and does very well in school." Children in the negative relevant condition heard "Casey is not a smart boy. He doesn't know very many things and he does poorly in school." Children were asked a peer performance question based on condition (e.g., in the positive, relevant condition they were asked "Did Casey get 11 stars because he is smart?"). Scoring and self-evaluations questions were the same as those used in Experiment 1 (Appendix S7). Reliability for self-evaluation explanations was high ( $\kappa = .89-.95$ ).

### Results

#### Peer Performance Question

*t*-Tests against chance indicated that older children only reliably endorsed "smart" as the cause of

the peer's performance (92%),  $p < .01$ . Younger children reliably endorsed "smart" (92%),  $p < .01$ , and "athletic" as the cause of the peer's performance marginally more often than expected by chance (75%),  $p = .08$ .

### Self-Evaluations

Self-evaluations were analyzed with a series of 2 (age: 5- to 6- year-olds vs. 9- to 10-year-olds)  $\times$  2 (trait valence: positive vs. negative)  $\times$  2 (trait relevance: relevant vs. irrelevant) between-subjects analyses of variance. There were no significant effects or interactions involving gender; thus, it was excluded from final models (see Table 1 for descriptive data).

*Affect.* There were significant main effects of age,  $F(1, 90) = 10.89$ ,  $p = .001$ ,  $\eta_p^2 = .11$ , and trait valence,  $F(1, 90) = 6.21$ ,  $p = .02$ ,  $\eta_p^2 = .07$ , that were qualified by a significant interaction between age and trait valence,  $F(1, 90) = 11.17$ ,  $p = .001$ ,  $\eta_p^2 = .11$ . Younger children's affect ratings were relatively high and did not differ significantly when the peer was described with a positive trait ( $M = 14.76$ ,  $SD = 4.24$ ) versus a negative trait ( $M = 15.32$ ,  $SD = 2.84$ ),  $t(47) = 0.55$ ,  $p = .59$ . In contrast, older children reported higher affect ratings when the peer was described with a positive trait ( $M = 14.79$ ,  $SD = 2.17$ ) as opposed to a negative trait ( $M = 10.87$ ,  $SD = 3.44$ ),  $t(46) = -4.725$ ,  $p < .001$ , irrespective of trait relevance (Figure 1b). Older children's affect in the positive condition did not differ significantly from younger children's affect ratings in the positive,  $t(36) = -0.03$ ,  $p = .97$ , or negative condition,  $t(44) = 0.73$ ,  $p = .47$ . No other main effects or interactions were significant,  $ps > .05$ .

*Performance.* Older children rated their performance lower than younger children,  $F(1, 90) = 8.52$ ,  $p = .004$ ,  $\eta_p^2 = .09$ . There was a significant interaction between age and trait valence,  $F(1, 90) = 3.94$ ,  $p = .05$ ,  $\eta_p^2 = .04$  (Figure 1c). Younger children's performance ratings were high and did not differ significantly when the peer was described with a positive trait as opposed to a negative trait,  $t(48) = 0.06$ ,  $p = .95$ . Older children rated their performance more favorably when the peer was described with a positive trait as compared with a negative trait,  $t(44) = -5.05$ ,  $p < .001$ . Older children's performance ratings in the positive condition did not differ significantly from younger children's performance ratings in the positive,  $t(33) = 0.70$ ,  $p = .48$ , or negative condition,  $t(32) = 0.78$ ,  $p = .44$ . No other main effects or interactions were significant,  $ps > .05$ .

*Ability.* Ability ratings decreased with age,  $F(1, 90) = 19.08$ ,  $p < .001$ ,  $\eta_p^2 = .18$ , but were relatively high (Table 1). No other main effects or interactions were significant,  $ps > .05$ .

### Affect, Performance, and Ability Explanations

See Appendix S8 for descriptive data and Appendix S9 for analyses.

### Entity/Incremental Reasoning

Correlational analyses indicated that ability (i.e., entity) explanations increased with age,  $r(96) = .28$ ,  $p = .007$ . Younger children provided at least one ability explanation 12% of the time, whereas older children did so 33.4% of the time.

### Discussion

As in Experiment 1, younger children reported high self-evaluations regardless of comparison peer characteristics. In contrast, older children's affective ratings and performance evaluations were lower when they were outperformed by peers with negative traits and higher when they were outperformed by peers with positive traits. By late childhood, peer traits clearly affect self-evaluation, suggesting that trait information alters children's interpretation of relative failure. This effect emerges later than children's ability to reason about their own and others' traits separately (Kurtz-Costes et al., 2005). Young children's disregard of relative failure may reflect views of ability as flexible (Nicholls, 1978). These children explained academic outcomes on the entity/incremental task in terms of malleable factors (e.g., effort). In part, they may not have understood the relevance of peer traits for performance, as is evident by their endorsement of athleticism as a cause of task performance. We extend previous work (Rhodes & Brickman, 2008) by establishing that peers' ability-related traits become relevant to self-evaluation later than more essentialized categories (e.g., gender; Taylor, Rhodes, & Gelman, 2009).

As expected, older children who were outperformed by peers with negative rather than positive traits reported less positive affect and lower performance ratings. As seen previously (Heyman & Dweck, 1998) and here, essentialist beliefs about intelligence may make peer traits seem more indicative of fixed ability and emphasize their importance for various self-evaluations. Surprisingly, peer traits were influential even when they were immaterial to

the task. Although older children view athletic progress as irrelevant to cognitive performance, provision of this information may have prompted them to assume that it was relevant. It is possible that older children exhibited a halo effect when reasoning about traits in a comparison context. That is, children may have viewed irrelevant traits as indicative of other similarly valenced traits. For example, they may not have viewed athleticism *per se* as the cause of a peer's performance, but inferred that an athletic peer is likely to be smart. This may explain why older children only endorsed "smart" as a cause of the peer's performance, but viewed irrelevant traits (i.e., athleticism) as meaningful to their own performance. As noted previously, even older children's trait reasoning can be limited in self-evaluative contexts (Schuster et al., 1998).

### General Discussion

We examined the impact of developmentally relevant peer comparison characteristics—expertise and ability-related traits—on self-evaluations in early to late childhood. Altogether, young children were impervious to comparative feedback, reporting positive affect and high performance evaluations regardless of peer characteristics. By late childhood, even modest failure relative to a peer (i.e., 2-point difference in scores) influenced self-evaluations. Further, these children adjusted their self-evaluation based on peers' ability-related traits and, in some cases, expertise. Across studies, ability evaluations were high despite relative failure. The studies uniquely document that with age, peer characteristics matter for children's responses to relative failure. Although there is currently no comprehensive developmental model of social comparison, ability conception theories provide a context for interpreting the findings.

Although ability conceptions purportedly underlie the development of social comparison sensitivity (e.g., Butler, 1989), little research has addressed their contribution directly. These studies are the first to provide support for ability conceptions as pivotal in the emergent effects of comparative feedback on self-evaluations. We propose that the provision of expertise, as compared with ability-related traits, likely cues reflection about different ability conceptions across early and middle childhood because of the differences in the extent to which these characteristics are essentialized. Specifically, expertise cues both younger and older children's conceptions of ability as malleable (Lockhart et al.,

2016), thereby limiting the effects of comparative feedback on self-evaluations (e.g., inconsistent and less robust effects in Experiment 1 as compared with Experiment 2). Conversely, trait information cues malleable ability reasoning in younger children and essentialized ability reasoning in older children, as these are the predominant ways in which children view ability-related traits at these ages (Gelman et al., 2007). This may explain why trait information resulted in more robust effects of comparative feedback on older, but not younger children's self-evaluations. Even traits do not cue essentialist reasoning as much as other social categories (e.g., gender) in early childhood (Rhodes & Brickman, 2008).

Three aspects of the current findings provide support for the interpretation above. First, both studies document the emergence of comparative feedback as influential to self-evaluations between early and middle childhood, the developmental period in which fixed ability conceptions emerge (Nicholls, 1978). Second, the available data indicate that children experienced this developmental shift in ability conceptions. Older children provided more ability explanations than younger children for real-life academic outcomes (i.e., Entity/Incremental task), and younger children rarely provided such explanations. Third, the provision of expertise as compared with ability-related traits had differential effects on children's self-evaluations and explanations. When expertise was provided, younger and older children explained their ability evaluations with explicit reference to ability (e.g., "I'm smart") as well as effort (e.g., "I do them all the time"). Conversely, when trait information was provided, younger children largely explained their ability evaluations in terms of effort, whereas older children largely referred to ability. These responses provide evidence that expertise likely cued younger and older children to reflect on the importance of practice. In contrast, the provision of ability-related traits cued older children to reflect on the importance of inherent skill. Developmental changes in positivity and reflection on fixed ability are likely fostered by parents' and teachers' emphasis on personal growth in early childhood and relative performance in middle childhood (Cimpian, 2017).

A major contribution of this research is its focus on a novel element of ability conceptions in the context of social comparisons, namely expertise. The fact that expertise and ability-related traits had some similar effects on children's self-evaluations (i.e., performance evaluations: familiar domain in Experiment 1 and results of Experiment 2) suggests

that these conceptions are related. Given the increasing essentialist beliefs about ability-related traits with age, children may begin to appreciate ability as a causal factor in achieving expertise. For example, high-ability individuals may be expected to attain expertise more quickly and with less effort than low-ability individuals. With age, children come to believe that high-ability individuals need less time and effort to achieve high performance (Nicholls & Miller, 1984). Similar conceptions may emerge for beliefs about ability as influential to the mastery of specialized knowledge. Early in life, these two concepts are not likely well differentiated, as effort is conceived of as the most important factor in obtaining either (Heyman, Gee, & Giles, 2003; Lockhart et al., 2016). Children's inferences about the relation between expertise and traits may be similar to their ideas about behavior-trait relations. Children more easily make trait-to-behavior inferences (e.g., a nice person will share) than behavior-to-trait inferences (e.g., a person who shares is nice; Liu, Gelman, & Wellman, 2007). The essentialist nature of traits may also make ability-to-expertise inferences (e.g., an artistic individual will know a lot about drawing) easier than expertise-to-ability inferences.

Our findings also suggest that social comparison is constrained by information processing demands. In contrast to traits, expertise may have had limited effects on self-evaluations because of informational complexity. Children must consider what it means to know "a lot" or "a little" and assess it in relation to their own knowledge. This context may impose working memory demands that limit children's ability to consider both their own and others' expertise, especially in a comparison context in which both are relevant. These effects may be particularly pronounced for younger children (see Morra, Parrella, & Camba, 2011). In contrast, trait information consisted of simplistic, familiar labels that map strongly onto behavioral outcomes (Kurtz-Costes et al., 2005). Thus, it may be easier for children to assess their own and others' traits as similar or different. Had we described expertise in a trait-like fashion, such as labeling professions (e.g., doctor; Lutz & Keil, 2002), the effects may have been similar to trait labels.

There are many fruitful directions for future research. It is important to understand which level of expertise (i.e., high vs. low) or trait type (i.e., positive vs. negative) accounts mainly for our findings. We propose that self-evaluations *improve* when peers are described with high expertise or positive traits (relative to unspecified peer

characteristics) given that self-evaluations of older children were as high as those in younger children. In contrast, when peer characteristics are not described, older children report lower self-evaluations than younger children after relative failure (Ruble et al., 1994). An unspecified peer condition is needed to determine whether failure relative to novices and peers with negative traits impairs self-evaluations further.

Further research is necessary to establish whether other peer characteristics may be more salient and thereby, more influential to younger children's use of comparative feedback. Future studies could use age as a proxy for denoting expertise or ability, as even preschoolers recognize age as an indicator of expertise (VanderBorghet & Jaswal, 2009). It may be easier for young children to use comparison peer age to inform self-evaluations (e.g., lower self-evaluations when outperformed by a younger child, but not an adult). This situation likely better reflects children's daily experiences than the direct provision of expertise information.

Future work should examine social comparison effects in multiple domains. Children may be more responsive to comparative feedback in familiar and significant domains (Bers & Rodin, 1984), which likely vary across development. For example, math and reading are content areas that are important to older children's daily lives, whereas activities like writing one's name may be more significant to young children. Relative failure in these areas may have greater effects on self-evaluations, including ability evaluations, which were unaffected here.

Given the ubiquity of social comparison (Frey & Ruble, 1985), these results have implications for children's daily lives. Awareness of relative failure can be detrimental, but it provides a chance for children to note areas of improvement. Thus, identification of situations in which children learn from negative feedback, rather than disregard it or respond helplessly, is important. Also, research on learning orientations (Kamins & Dweck, 1999) may be informative for understanding individual differences in social comparison and self-evaluation in childhood.

## References

- Aguiar, N. R., Stoess, C. J., & Taylor, M. (2012). The development of children's ability to fill the gaps in their knowledge by consulting experts. *Child Development, 83*, 1368-1381. <https://doi.org/10.1111/j.1467-8624.2012.01782.x>
- Altermatt, E. R., & Pomerantz, E. M. (2003). The development of competence-related and motivational beliefs:

- An investigation of similarity and influence among friends. *Journal of Educational Psychology*, 95, 111–123. <https://doi.org/10.1037/0022-0663.95.1.111>
- Benenson, J. F., & Dweck, C. S. (1986). The development of trait explanations and self-evaluations in the academic and social domains. *Child Development*, 57, 1179–1187. <https://doi.org/10.2307/1130441>
- Bers, S. A., & Rodin, J. (1984). Social-comparison jealousy: A developmental and motivational study. *Journal of Personality and Social Psychology*, 47, 766–779. <https://doi.org/10.1037/0022-3514.47.4.766>
- Boseovski, J. J. (2010). Evidence for “rose-colored glasses”: An examination of the positivity bias in young children’s personality judgments. *Child Development Perspectives*, 4, 212–218. <https://doi.org/10.1111/j.1750-8606.2010.00149.x>
- Boseovski, J. J., Hughes, C., & Miller, S. E. (2016). Expertise in unexpected places: Children’s acceptance of information from gender counter-stereotypical experts. *Journal of Experimental Child Psychology*, 141, 161–176. <https://doi.org/10.1016/j.jecp.2015.09.002>
- Boseovski, J. J., & Thurman, S. L. (2014). Evaluating and approaching a strange animal: Children’s trust in informant testimony. *Child Development*, 85, 824–834. <https://doi.org/10.1111/cdev.12156>
- Butler, R. (1989). Interest in the task and interest in peers’ work in competitive and noncompetitive conditions: A developmental study. *Child Development*, 60, 562–570. <https://doi.org/10.2307/1130722>
- Butler, R. (1992). What young people want to know when: Effects of mastery and ability goals on interest in different kinds of social comparisons. *Journal of Personality and Social Psychology*, 62, 934–943. <https://doi.org/10.1037/0022-3514.62.6.934>
- Cimpian, A. (2010). The impact of generic language about ability on children’s achievement motivation. *Developmental Psychology*, 46, 1333–1340. <https://doi.org/10.1037/a0019665>
- Cimpian, A. (2017). Early reasoning about competence is not irrationally optimistic, nor does it stem from inadequate cognitive representations. In A. J. Elliot, C. S. Dweck & D. S. Yeager (Eds.), *Handbook of competence and motivation: Theory and application* (2nd ed., pp. 387–407). New York, NY: Guilford Press.
- Danovitch, J. H., & Keil, F. C. (2007). Choosing between hearts and minds: Children’s understanding of moral advisors. *Cognitive Development*, 22, 110–123. <https://doi.org/10.1016/j.cogdev.2006.07.001>
- Frazier, B. N., & Gelman, S. A. (2009). Developmental changes in judgments of authentic objects. *Cognitive Development*, 24, 284–292. <https://doi.org/10.1016/j.cogdev.2009.06.003>
- Frey, K. S., & Ruble, D. N. (1985). What children say when the teacher is not around: Conflicting goals in social comparison and performance assessment in the classroom. *Journal of Personality and Social Psychology*, 48, 550–562. <https://doi.org/10.1037/0022-3514.48.3.550>
- Gelman, S. A., Heyman, G. D., & Legare, C. H. (2007). Developmental changes in the coherence of essentialist beliefs about psychological characteristics. *Child Development*, 78, 757–774. <https://doi.org/10.1111/j.1467-8624.2007.01031.x>
- Hermes, J., Behne, T., & Rakoczy, H. (2015). The role of trait reasoning in young children’s selective trust. *Developmental Psychology*, 51, 1574–1587. <https://doi.org/10.1037/dev0000042>
- Heyman, G. D., & Dweck, C. S. (1998). Children’s thinking about traits: Implications for judgments of the self and others. *Child Development*, 69, 391–403. <https://doi.org/10.2307/1132173>
- 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., & Gelman, S. A. (2000). Preschool children’s use of trait labels to make inductive inferences. *Journal of Experimental Child Psychology*, 77, 1–19. <https://doi.org/10.1006/jecp.1999.2555>
- Heyman, G. D., & Giles, J. W. (2004). Valence effects in reasoning about evaluative traits. *Merrill-Palmer Quarterly*, 50, 86–109. <https://doi.org/10.1353/mpq.2004.0004>
- 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. <https://doi.org/10.1037/0012-1649.35.3.835>
- Koenig, M. A., & Jaswal, V. K. (2011). Characterizing children’s expectations about expertise and incompetence: Halo or pitchfork effects? *Child Development*, 82, 1634–1647. <https://doi.org/10.1111/j.1467-8624.2011.01618.x>
- Kurtz-Costes, B., McCall, R. J., Kinlaw, C. R., Wiesen, C. A., & Joyner, M. H. (2005). What does it mean to be smart? The development of children’s beliefs about intelligence in Germany and the United States. *Journal of Applied Developmental Psychology*, 26, 217–233. <https://doi.org/10.1016/j.appdev.2004.12.005>
- Landrum, A. R., Mills, C. M., & Johnston, A. M. (2013). When do children trust the expert? Benevolence information influences children’s trust more than expertise. *Developmental Science*, 16, 622–638. <https://doi.org/10.1111/desc.12059>
- Liu, D., Gelman, S. A., & Wellman, H. M. (2007). Components of young children’s trait understanding: Behavior-to-trait inferences and trait-to-behavior predictions. *Child Development*, 78, 1543–1558. <https://doi.org/10.1111/j.1467-8624.2007.01082.x>
- Lockhart, K. L., Goddu, M. K., & Keil, F. C. (2016). Overoptimism about future knowledge: Early arrogance? *Journal of Positive Psychology*, 87, 477–493. <https://doi.org/10.1080/17439760.2016.1167939>
- Lockhart, K. L., Keil, F. C., & Aw, J. (2013). A bias for the natural? Children’s beliefs about traits acquired through effort, bribes, or medicine. *Developmental Psychology*, 49, 1669–1682. <https://doi.org/10.1037/a0030769>
- Lutz, D. J., & Keil, F. C. (2002). Early understanding of the division of cognitive labor. *Child Development*, 73, 1073–1084. <https://doi.org/10.1111/1467-8624.00458>

- Morra, S., Parrella, I., & Camba, R. (2011). The role of working memory in the development of emotion comprehension. *British Journal of Developmental Psychology, 29*, 744–764. <https://doi.org/10.1348/2044-835X.002006>
- Morris, W. N., & Nemcek, D. (1982). The development of social comparison motivation among preschoolers: Evidence of a stepwise progression. *Merrill-Palmer Quarterly, 28*, 413–425.
- 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. <https://doi.org/10.2307/1128250>
- Nicholls, J. G., & Miller, A. T. (1984). Reasoning about the ability of self and others: A developmental study. *Child Development, 55*, 1990–1999. <https://doi.org/10.2307/1129774>
- Pomerantz, E. M., Ruble, D. N., Frey, K. S., & Greulich, F. (1995). Meeting goals and confronting conflict: Children's changing perceptions of social comparison. *Child Development, 66*, 723–738. <https://doi.org/10.2307/1131946>
- Raven, J. C., Court, J. H., & Raven, J. (1995). *Coloured progressive matrices*. Oxford, UK: Oxford Psychologists.
- Rhodes, M., & Brickman, D. (2008). Preschoolers' responses to social comparisons involving relative failure. *Psychological Science, 19*, 968–972. <https://doi.org/10.1111/j.1467-9280.2008.02184.x>
- Rhodes, M., & Gelman, S. A. (2008). Categories influence predictions about individual consistency. *Child Development, 79*, 1270–1287. <https://doi.org/10.1111/j.1467-8624.2008.01188.x>
- Ruble, D. N., Boggiano, A. K., Feldman, N. S., & Loebel, J. H. (1980). Developmental analysis of the role of social comparison in self-evaluation. *Developmental Psychology, 16*, 105–115. <https://doi.org/10.1037/0012-1649.16.2.105>
- Ruble, D. N., Eisenberg, R., & Higgins, E. T. (1994). Developmental changes in achievement evaluation: Motivational implications of self-other differences. *Child Development, 65*, 1095–1110. <https://doi.org/10.2307/1131307>
- Ruble, D. N., & Flett, G. L. (1988). Conflicting goals in self-evaluative information seeking: Developmental and ability level analyses. *Child Development, 59*, 97–106. <https://doi.org/10.2307/1130392>
- Schunk, D. H., & Pajares, F. (2002). The development of academic self-efficacy. In A. Wigfield & J. S. Eccles (Eds.), *Development of achievement motivation* (pp. 15–31). San Diego, CA: Academic Press.
- 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/10.2307/1132133>
- Taylor, M. G., Rhodes, M., & Gelman, S. A. (2009). Boys will be boys; cows will be cows: Children's essentialist reasoning about gender categories and animal species. *Child Development, 80*, 461–481. <https://doi.org/10.1111/j.1467-8624.2009.01272.x>
- VanderBorghet, M., & Jaswal, V. K. (2009). Who knows best? Preschoolers sometimes prefer child informants over adult informants. *Infant and Child Development, 18*, 61–71. <https://doi.org/10.1002/icd.591>
- Welsh, M. C. (1991). Rule-guided behavior and self-monitoring on the Tower of Hanoi disk-transfer task. *Cognitive Development, 6*, 59–76. [https://doi.org/10.1016/0885-2014\(91\)90006-Y](https://doi.org/10.1016/0885-2014(91)90006-Y)

### Supporting Information

Additional supporting information may be found in the online version of this article at the publisher's website:

**Appendix S1.** Experiment 1 Peer Expertise Information

**Appendix S2.** Experiment 1 Self-Evaluation Explanation Coding and Examples

**Appendix S3.** Experiment 1 Descriptive Data for Self-Evaluation Explanations

**Appendix S4.** Experiment 1 Analyses of Age-Related Change in Self-Evaluation Explanations

**Appendix S5.** Experiment 2 Entity/Incremental Questions and Coding

**Appendix S6.** Experiment 2 Peer Trait Information

**Appendix S7.** Experiment 2 Self-Evaluation Explanation Coding

**Appendix S8.** Experiment 2 Descriptive Data for Self-Evaluation Explanations

**Appendix S9.** Experiment 2 Analyses of Age-Related Change in Self-Evaluation Explanations

**Appendix S10.** Comprehension Questions from Experiments 1 and 2