

# Selective Skepticism: American and Chinese Children's Reasoning About Evaluative Academic Feedback

Gail D. Heyman  
University of California, San Diego

Genyue Fu  
Zhejiang Normal University

Kang Lee  
University of Toronto

Children's reasoning about the credibility of positive and negative evaluations of academic performance was examined. Across 2 studies, 7- and 10-year-olds from the United States and China ( $N = 334$ ) judged the credibility of academic evaluations that were directed toward an unfamiliar peer. In Study 1, participants from China responded that criticism should be accepted to a greater extent than did participants from the United States, and children from both countries demonstrated a selective skepticism effect by treating negative feedback more skeptically than positive feedback. Study 2 replicated the selective skepticism effect among children from both countries and ruled out the possibility that it can be explained as a rational analysis of perceived base rates. The results suggest that children are selective in their trust of evaluative feedback and that their credibility judgments may be influenced by the desirability of the information that is being conveyed or its anticipated consequences.

*Keywords:* evaluation, feedback, reasoning, cross-cultural differences, social cognition

If children are to make effective use of the information they obtain from others, it is important that they learn to evaluate it critically. Recent studies have documented that preschool-age children make these evaluations using a variety of strategies (Gelman, 2009; Harris, 2007; Heyman, 2008), such as considering an informant's accuracy on previous occasions (Birch, Vauthier, & Bloom, 2008; Corriveau & Harris, 2009; Jaswal & Neely, 2006; Koenig & Harris, 2005), the level of confidence that the informant has expressed (Sabbagh & Baldwin, 2001), or the informant's perceptual access to relevant information (Robinson, Champion, & Mitchell, 1999). However, children sometimes have difficulty critically evaluating what they are told (Moses & Baldwin, 2005). For example, 3- and 4-year-olds may fail to make use of information that indicates deceptive intent (Lee & Cameron, 2000; Vanderbilt, Liu, & Heyman, 2011), and children younger than about age 8 have trouble anticipating the potential effects of motives relating to self-interest or biases due to personal relationships (Heyman & Legare, 2005; Mills & Grant, 2009; Mills & Keil, 2005, 2008).

One factor that may serve to constrain children's ability to critically evaluate what they learn from others is the desirability of the information being conveyed. This possibility is consistent with recent evidence that throughout middle childhood, children show a tendency to evaluate information about themselves and others in positive ways (see Boseovski, 2010). For example, they generally require less behavioral evidence to make positive trait attributions than to make negative trait attributions. There is also evidence that children's desires can influence the conclusions they reach (Woolley, Boerger, & Markman, 2004). However, it is not clear whether children are predisposed to be unduly skeptical about undesirable evaluations when they have no personal stake in the outcome.

Our primary goal in this paper was to make a direct test of the possibility that children will show greater skepticism of evaluative information when it is negative than when it is positive, which we refer to as the *selective skepticism hypothesis*. We asked participants to reason about the credibility of evaluative feedback that was directed toward others. We focused on evaluative feedback because it provides a means to examine the selective skepticism hypothesis and also because of its implications for how children come to view themselves, including how they develop a sense of their own capabilities (Altermatt, Pomerantz, Ruble, Frey, & Greulich, 2002; Dweck, 1999; Parsons, Kaczala, & Meece, 1982). For example, it is possible that children who are overly receptive to positive academic feedback but highly skeptical of negative feedback will have only a limited understanding of their own shortcomings and how such shortcomings can be remedied.

We examined these issues from a cross-cultural perspective to assess the extent to which selective skepticism might reflect culturally specific processes. We compared children from the United States (U.S.) and China because of differences in cultural values about how evaluative feedback should be interpreted. In China,

---

This article was published Online First December 31, 2012.

Gail D. Heyman, Department of Psychology, University of California, San Diego; Genyue Fu, School of Education, Zhejiang Normal University, Jinhua, Zhejiang Province, People's Republic of China; Kang Lee, Ontario Institute for Studies in Education, University of Toronto, Toronto, Ontario, Canada.

This research was supported by Grant GDW20123300141 from the High End Foreign Experts Recruitment Program of China.

Correspondence concerning this article should be addressed to Gail D. Heyman, Department of Psychology, University of California, San Diego, 9500 Gilman Drive, La Jolla, CA 92093-0109. E-mail: gheyman@ucsd.edu

there is less of a concern that negative feedback will harm self-esteem and more of an emphasis on the potential of negative feedback to promote better performance in the future (Hess, Chang, & McDevitt, 1987; Heyman, Fu, Sweet, & Lee, 2009; Miller, Wang, Sandel, & Cho, 2002; Ng, Pomerantz, & Lam, 2007; Stigler & Perry, 1990). Ng et al. (2007) asked fifth graders from the U.S. and from China about their parents' responses to their successful and unsuccessful academic performance. Chinese participants reported that their parents tended to emphasize their failures and deemphasize their successes, and participants from the U.S. reported the opposite pattern. Ng et al. replicated this general result in a laboratory task in which parents responded to the success or failure of their fourth- and fifth-grader children on a problem-solving task. Heyman et al. (2009) found that 10- and 11-year-olds in China often mentioned the concern that hearing positive feedback could make an individual become "too proud," but that this concern was not mentioned by children of the same age in the U.S.

These cross-cultural differences appear to extend to the way teachers evaluate children's performance in the classroom. Stigler and Perry (1990) found that when teaching math, first- and fifth-grade teachers in the U.S. gave more positive evaluative feedback than did their Chinese counterparts. There is also evidence that Chinese teachers tend to see criticism as a helpful way to encourage children to recognize shortcomings and strive to overcome them but are wary of offering children too much praise because they see it as inconsistent with promoting modesty and motivational resilience (Wang, Liu, & Yin, 2009). Moreover, beginning as soon as early elementary school, teachers in China often announce test results to the class and publicly identify the worst performers as a means to motivate them to improve and to learn from their more successful peers. It may be that these teaching practices are derived from broader beliefs about the nature of learning within Chinese culture, in which a primary goal of learning is to perfect oneself socially and morally, and maintaining humility and experiencing emotions such as shame and guilt are seen as functioning to motivate individuals to reach this goal (Li, 2005).

This evidence of cross-cultural differences in the way children reason about positive versus negative feedback led us to predict that children in the U.S. would tend to accept positive feedback and discount negative feedback to a greater extent than would children in China. We included samples from both urban and rural settings in China to allow for an examination of within-country variability, because individuals from rural backgrounds are more likely to be socialized in a way that emphasizes traditional Chinese values such as being group oriented (Chen, Wang, & Wang, 2009; Zhang & Fuligni, 2006). It is likely that the large-scale social and economic changes that have taken place in China over the last quarter century, including a transition to a market economy that tends to reward individual efforts, have been felt more strongly by children in urban settings than in rural settings, which tend to be agricultural and tend to offer limited opportunities for travel (Zhang & Fuligni, 2006). Recent research has documented that this difference between rural and urban environments has meaningful psychological consequences (Chen et al., 2009; Fu et al., 2010; Helwig, Yang, Tan, Liu, & Shao, 2011; Zhang & Fuligni, 2006). For example, within the context of traditional Chinese values, the characteristic of shyness is often viewed favorably, which con-

trasts with the negative view that is generally held in North America (Chen, Cen, Li, & He, 2005; Chen et al., 2009). Chen et al. (2009) found that among urban Chinese residents, shyness was associated with social and school problems, but that among rural migrants it was associated with indexes of being well adjusted. Zhang and Fuligni (2006) found that as compared to rural Chinese adolescents, urban Chinese adolescents tended to have more intense conflict with their parents. These differences led us to predict that children from a rural Chinese background would be even less likely than children from urban communities to selectively accept positive feedback over negative feedback.

## The Present Research

In our evaluation of the selective skepticism hypothesis, we asked children to reason about feedback that was directed toward an unfamiliar individual. The purpose of making the target an unfamiliar individual was to maximize experimental control, by ensuring that children would have no prior knowledge about the ability or past performance of the person being evaluated. In addition, if children were to show a selective skepticism effect with reference to information about themselves or someone they know it would be difficult to determine whether the effect is limited to cases in which children have a direct stake in the outcome.

Participants were 7- and 10-year-olds, an age range that was chosen in light of recent evidence of broader age-related changes in how children make psychological inferences, including inferences about self-presentational processes (Aloise-Young, 1993; Banerjee, 2000; Banerjee & Yuill, 1999; Benne & Yeeles, 1990; Heyman, Fu, & Lee, 2007; Watling & Banerjee, 2007). There is also evidence of increasing age-related sensitivity to cues that relate to possible biases involving self-interest or personal relationships (Heyman & Legare, 2005; Mills & Grant, 2009; Mills & Keil, 2005, 2008). For example, Heyman and Legare (2005) found that 6- and 7-year-olds tended to agree with the notion that one can discover whether an individual is truthful simply by asking the individual in question, and they often explained their responses in terms of what people should do (e.g., "If you ask them to tell the truth, then they better tell the truth, otherwise it would be a lie"). In contrast, 10- and 11-year-olds tended to reject this notion and often explained their responses with reference to the fact that people do not always accurately convey what they know (e.g., "You cannot always trust people by what they say because sometimes people make up what they talk about"). There is also evidence of changes in reasoning about evaluative messages during these ages, at least when reasoning about peer feedback: Heyman et al. (2009) found that in both the U.S. and China, younger children were more likely than older children to expect others to preferentially give peers positive evaluative feedback over negative evaluative feedback.

## Study 1

Participants from the U.S. and China were presented with a series of scenarios in which they were asked to judge the credibility of evaluations of academic work. The primary goal was to evaluate the selective skepticism hypothesis. We presented evaluations from both teachers and peers to determine the extent to

which children's reasoning would generalize across these two types of evaluators and also because both groups tend to serve as significant influences on children (Altermatt et al., 2002; Parsons et al., 1982).

In addition, we examined whether feedback that is directed toward all of the individuals in a class would be discounted to a greater extent than feedback that is directed toward a single individual. This can be viewed as a manipulation of distinctiveness, with the feedback directed to all individuals reflecting low distinctiveness and the feedback directed to one individual reflecting high distinctiveness (Boseovski & Lee, 2006; Kelley, 1973). We reasoned that in this context, low distinctiveness might imply a tendency on the part of the teacher to be critical or supportive in general, rather than simply evaluating the quality of a particular student's work on a single occasion.

We predicted that children in the U.S. would show a pattern of selective skepticism, but we did not expect to see such a pattern among the children in China, especially in rural China, where influences from the self-esteem culture that is seen in the U.S. are likely to be minimal. We also predicted that selective skepticism would decrease with age, based on evidence of developmental changes across the elementary school years. These changes include a decrease in positivity biases (see Boseovski, 2010) and an increase in cognitive skills such as the ability to reflect upon the implications of social information (Schuster, Ruble, & Weinert, 1998).

## Method

**Participants.** Participants were a total of 206 elementary school students from the U.S. and China. There were 74 participants from the U.S., 35 in a younger group ( $M = 7$  years 7 months, range = 7 years 0 months to 8 years 3 months, 15 girls) and 39 in an older group ( $M = 10$  years 11 months, range = 10 years 4 months to 11 years 4 months, 16 girls) who lived in a southwestern coastal city and attended schools whose students were primarily from middle-class and upper-middle-class backgrounds. The sample from the U.S. was approximately 85% Caucasian, 8% Asian American, 4% Hispanic American, and 3% African American.

There were 66 participants from urban China and 66 from rural China. At each site, 33 children were in a younger group ( $M = 7$  years 7 months, range = 7 years 3 months to 8 years 1 month, 16 girls in urban China;  $M = 7$  years 9 months, range = 7 years 4 months to 8 years 1 month, 18 girls in rural China), and 33 were in an older group ( $M = 10$  years 9 months, range = 10 years 4 months to 11 years 1 month, 16 girls in urban China;  $M = 10$  years 5 months, range = 10 years 1 month to 10 years 9 months, 16 girls in rural China). The urban Chinese participants attended schools whose students were primarily from a middle-class background, and the rural Chinese participants attended schools in a low-income rural community. All of the Chinese participants were Han Chinese.

**Procedure.** In individual interview sessions, participants were presented with a series of eight scenarios. The scenarios portrayed a student or group of students receiving evaluative feedback about essays that described a class field trip. In a  $2 \times 2 \times 2$  within-subjects manipulation, the scenarios differed in the valence of the evaluation (positive vs. negative), the source (teacher vs. peer), and the recipient (a single student or all individuals in the class).

Questions were asked in a random order that was determined separately for each participant, and participants either heard all scenarios with male characters or heard all scenarios with female characters. The following is an example of a scenario in which the positive feedback was provided by a teacher to an individual student (see also Appendix A).

Luke wrote an essay about a class field trip. His teacher, Mr. Adams, read Luke's essay before reading anyone else's and told Luke that his essay is very good. Should Luke believe that his essay is very good?

Children provided an initial yes or no response, and they then were asked a follow-up question to assess their certainty. For example, a child who responded "no" would be asked, "maybe no or really no?" The final response options were *really no* (scored as 0), *maybe no* (scored as 1), *maybe yes* (scored as 2), and *really yes* (scored as 3), which was used to create a *credibility rating* that ranged from 0 to 3. *Evaluative difference scores* were computed by subtracting the credibility ratings for negative feedback from the credibility ratings for positive feedback.

The questions were initially written in English and then translated into Chinese for the Chinese participants. To ensure that the translation process did not introduce changes in meaning, a translator who had not seen the original English versions translated the Chinese versions back into English. The back-translated versions of the stimuli did not differ in meaning from the original versions.

## Results

Mean credibility ratings are presented in Figure 1. Preliminary analyses showed no effects of participant gender, character gender, or recipient (i.e., a single student vs. all individuals in the class), so these variables were not included in subsequent analyses.

**Teacher credibility ratings.** A  $2$  (valence: positive, negative)  $\times 2$  (age group: younger, older)  $\times 3$  (location: U.S., urban China, rural China) analysis of variance (ANOVA) was conducted on the mean credibility ratings of teacher feedback. Valence was a within-subjects factor, and age group and location were between-subjects factors. There was a main effect of valence,  $F(1, 199) = 187.47, p < .001$ , partial  $\eta^2 = 1.00$ , with substantially greater acceptance of positive feedback ( $M = 2.39, SD = 0.61$ ) than negative feedback ( $M = 1.57, SD = 0.80$ ). There was a main effect of location,  $F(1, 199) = 7.65, p < .001$ , partial  $\eta^2 = 0.94$ , with slightly lower credibility ratings among children in the U.S. ( $M = 1.80, SD = 0.46$ ) than in urban China ( $M = 2.08, SD = 0.51$ ) or rural China ( $M = 2.07, SD = 0.55$ ). There was also a main effect of age group,  $F(1, 199) = 4.00, p < .05$ , partial  $\eta^2 = 0.51$ , with slightly higher credibility ratings among older children ( $M = 2.05, SD = 0.47$ ) than younger children ( $M = 1.91, SD = 0.57$ ).

There were three significant interactions. First, there was an interaction between valence and location,  $F(2, 199) = 18.28, p < .001$ , partial  $\eta^2 = 1.00$ , with larger valence effects in the U.S. than in urban or rural China: Evaluative difference scores were 1.29 ( $SD = 0.98$ ) in the U.S. versus 0.58 ( $SD = 0.81$ ) in urban China and 0.55 ( $SD = 0.85$ ) in rural China. Second, there was an interaction between valence and age group, with younger children reporting larger valence effects,  $F(1, 199) = 4.86, p < .05$ , partial  $\eta^2 = 0.59$ : Evaluative difference scores were 0.97 ( $SD = 1.03$ ) for younger children and 0.70 ( $SD = 0.85$ ) for older children. Finally, there was a significant three-

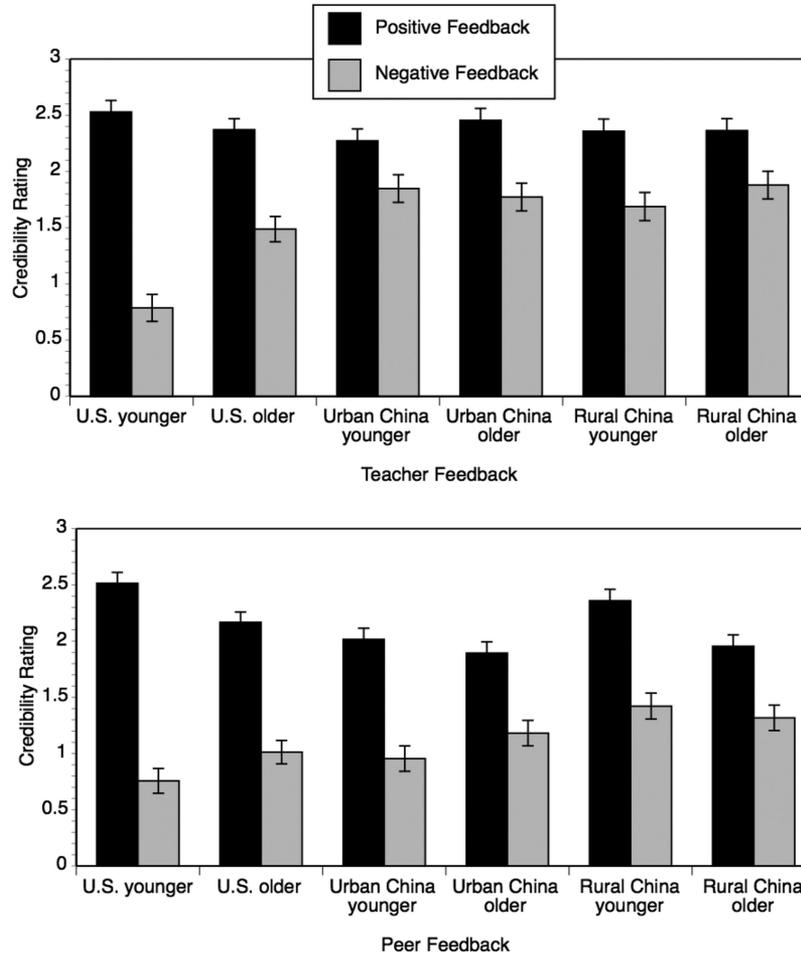


Figure 1. Mean credibility ratings of teacher and peer feedback for each group in Study 1, for both positive and negative feedback. Error bars indicate standard errors.

way interaction among valence, location, and age group, in which the age differences in valence effects were stronger in the U.S. than in urban or rural China. This result can be seen by subtracting the evaluative scores of older children from those of the younger children. In the U.S., the difference was 0.86 ( $SD = 0.89$ ), which indicates that there was a substantially stronger effect of valence among younger children than older children. In urban China this difference was 0.19 ( $SD = 0.81$ ), which suggests a weaker pattern in the same direction, and in rural China it was  $-0.41$  ( $SD = 0.84$ ), which suggests that there were somewhat stronger valence effects among the older children than among the younger children.

Follow-up analyses indicated that valence effects were significant within each location ( $p < .001$  within each group). Additional follow-up analyses that examined credibility ratings of praise versus criticism separately showed that there was no significant effect of location for praise, but there was for criticism. Least significant difference (LSD) tests indicated that participants in urban and rural China gave higher credibility ratings to criticism than did participants from the U.S. ( $M = 1.81$ ,  $SD = .72$  for urban China,  $1.78$ ,  $SD = 0.73$  for rural China, and  $1.16$ ,  $SD = 0.75$  for the U.S.).

**Peer credibility ratings.** A 2 (valence: positive, negative)  $\times$  2 (age group: younger, older)  $\times$  3 (location: U.S., urban China, rural China) ANOVA was conducted on mean credibility ratings of peer feedback. There was a main effect of valence:  $F(1, 199) = 276.70$ ,  $p < .001$ , partial  $\eta^2 = 1.00$ , with substantially greater acceptance of positive feedback ( $M = 2.15$ ,  $SD = 0.61$ ) than negative feedback ( $M = 1.11$ ,  $SD = 0.68$ ). There was also a main effect of location,  $F(1, 199) = 6.00$ ,  $p < .01$ , partial  $\eta^2 = 0.88$ , with slightly greater skepticism from children in the rural China ( $M = 1.76$ ,  $SD = 0.44$ ) than in the U.S. ( $M = 1.61$ ,  $SD = 0.39$ ) or in urban China ( $M = 1.51$ ,  $SD = 0.44$ ).

There was an interaction between valence and location,  $F(2, 199) = 11.45$ ,  $p < .001$ , partial  $\eta^2 = 0.993$ , with stronger valence effects in the U.S. than in urban or rural China (evaluative difference scores were  $1.45$ ,  $SD = 0.97$  for the U.S.;  $0.89$ ,  $SD = 0.89$  for urban China; and  $0.79$ ,  $SD = 0.89$  for rural China). There was also an interaction between age group and valence, with younger children showing greater selective skepticism than older children,  $F(1, 199) = 11.09$ ,  $p < .001$ , partial  $\eta^2 = .912$  (evaluative difference scores were  $1.25$ ,  $SD = 1.00$  for younger children, and  $0.83$ ,  $SD = 0.87$  for older children).

Follow-up analyses indicated that valence effects were significant in each location ( $p < .001$  within each group). Additional follow-up analyses examining credibility ratings of praise versus criticism separately indicated a willingness to accept both praise ( $p < .001$ ) and criticism ( $p < .001$ ). Regarding the acceptance of praise, LSD tests indicated that the only difference was that participants in urban China ( $M = 1.96$ ,  $SD = 0.58$ ) gave lower credibility ratings than did participants from the U.S. ( $M = 2.33$ ,  $SD = 0.54$ ). In evaluating criticism, LSD tests indicated that participants in rural China ( $M = 1.37$ ,  $SD = 0.66$ ) were less skeptical than those in the U.S. ( $M = 0.89$ ,  $SD = 0.69$ ) or urban China ( $M = 1.07$ ,  $SD = 0.61$ ).

## Discussion

The hypothesis that children would be more skeptical of evaluative feedback when it is negative rather than positive was strongly supported with reference to feedback from both teachers and peers. This selective skepticism effect was evident in the U.S., as predicted, but contrary to our predictions, was also seen in both rural and urban China.

We saw no evidence that children considered feedback that was directed toward all members of a class (low distinctiveness) to be less credible than feedback directed toward a single individual (high distinctiveness). This finding was also contrary to what we had predicted, based on the assumption that low distinctiveness might promote dispositional attributions about the teacher's tendency to be critical or supportive in general rather than to simply evaluate the quality of the work. Children's comments suggested that some did indeed make such an inference (e.g., several noted that there is no way all that students in the class would perform poorly, so the teacher should not be believed). However, children's comments also suggested that feedback was sometimes discounted because it was directed toward one individual. In these cases, children appeared to reason that the feedback reflected the teacher's evaluation of the student in general rather than the quality of the student's work on that specific occasion.

There was evidence of age-related and cross-cultural differences in children's judgments. The strongest selective skepticism effect was seen among the younger children from the U.S., who tended to assert that positive feedback should be accepted and negative feedback should be rejected regardless of whether it comes from teachers or peers. The other groups of participants were still highly accepting of positive feedback from teachers, but they did not clearly reject negative feedback from teachers, and they tended to be more skeptical of feedback when it came from peers.

Notably, the pattern of reasoning among the rural Chinese participants was very similar to that of the urban Chinese participants. This is somewhat surprising in light of evidence of certain psychological consequences to growing up in these different contexts (Chen et al., 2009; Fu et al., 2010; Helwig et al., 2011; Zhang & Fuligni, 2006). It may be that because such urban-rural differences tend to be closely tied to traditional Chinese values about how to function in a group context, they are not particularly relevant to the question of how individuals reason about evaluative feedback. Future studies could test this issue by contrasting evaluative feedback given to a group versus an individual.

## Study 2

As predicted, Study 1 provided evidence of selectivity in children's skepticism. Although this pattern was strongest among children from the U.S., it was also seen among children from rural and urban China on all measures. In considering these results, it is reasonable to ask whether a pattern of selective skepticism might constitute a rational analysis of the available evidence. If one were to assume that the base rate of work that is "very good" is greater than the base rate of work that is "not good," selective skepticism might be a rational approach. To evaluate this alternative explanation for the results of Study 1, we modified the feedback to indicate that the teacher describes a student's work as either the best in the class or the worst in the class, based on the assumption that for a given assignment there will only be one student whose performance is the very best and one student whose performance is the very worst.

It is also possible that the pattern of selective skepticism seen in Study 1 was a function of children's expectations about the base rate of false communication within the domain. For example, if one were evaluating self-reported information about body weight, extra skepticism would be warranted if the information has been obtained from a dating-related website for which users are known to misrepresent their weight. Similarly, selective skepticism might be rational in contexts in which evaluators are known to provide more false positive feedback than false negative feedback. To examine this possibility, we assessed children's beliefs about the perceived likelihood of false positive and false negative feedback.

Because the results of Study 1 were similar regardless of whether (a) the Chinese participants were urban or rural, (b) the feedback was directed toward a single student or the entire class, or (c) the person offering the feedback was a teacher or a peer, these three contrasts were omitted from Study 2. The Chinese participants in Study 2 were from an urban location only. No information was provided about whether or not other individuals had also heard the feedback. We chose to present feedback from a teacher only because there are more reasons to legitimately question the credibility of peer feedback regarding academic matters (e.g., peers might lack the necessary expertise on academic matters to provide meaningful feedback), which made findings relating to peer feedback more difficult to interpret. A final change to the procedure of Study 2 was that we added a between-subjects manipulation of the subject area (writing vs. science) to examine whether it would affect children's credibility judgments.

## Method

**Participants.** Participants were a total of 128 elementary school students from the U.S. and China. There were 59 participants from the U.S., 32 in a younger group ( $M = 7$  years 7 months, range = 7 years 0 months to 8 years 1 month, 14 girls) and 27 in an older group ( $M = 10$  years 10 months, range = 10 years 3 months to 11 years 10 months, 12 girls). The U.S. sample was approximately 69% Caucasian, 24% Asian American, 2% Hispanic American, and 5% African American and was similar demographically to the U.S. sample from Study 1.

There were a total of 66 participants from China, 33 in a younger group ( $M = 7$  years 3 months, range = 6 years 10 months to 7 years 7 months, 17 girls) and 33 in an older group ( $M = 10$  years 3 months, range = 9 years 7 months to 11 years 6 months,

22 girls). The sample was 100% Han Chinese and was similar demographically to the urban Chinese sample from Study 1.

**Procedure.** In individual interviews, participants were presented with two pairs of questions. The first pair involved asking children to make credibility ratings of two scenarios in which a student receives evaluative feedback from a teacher. The valence of the feedback (positive vs. negative) was manipulated as a within-subjects factor, and the order of the two questions was randomized for each participant. The subject area of the feedback (writing or science) was manipulated as a between-subjects factor, as was the gender of the character. Feedback was made more specific than in Study 1 by indicating that the student was either the best or the worst performer in the class. Because a randomly chosen individual is equally likely to be the best or the worst in the class at a particular subject, this wording controls for base rates of positive and negative performance. The following is a scenario that involved writing and positive feedback (also see appendix A).

Christopher wrote an essay for his homework. His teacher, Mr. Turner, told Christopher that his essay was the worst in the class. Should Christopher believe Mr. Turner that his essay was the worst in the class?

The response scales concerning teacher feedback were the same as in Study 1, as were the translation and back-translation procedure. However, participants in Study 2 were asked to explain their answers. Most of the responses contained at least one reference to the question of whether teachers should be viewed as good sources of information, the truth value of the information that was conveyed, advice to the recipient of the information, factors having to do with the desirability or motivational implications of the feedback, or some combination.

Of theoretical interest were responses that made reference to factors having to do with the desirability or motivational implications of the feedback. Examples of responses that fell into this category and responses that did not are presented in Appendix B. When these concepts were mentioned, they were further coded with reference to the notion that it is better to believe that one has high ability (e.g., “he should believe in himself,” “believe the praise because it is a good thing,” and “if he believes that his work isn’t good, he will lose confidence”) versus the notion that it can be problematic to believe that one has high ability (e.g., “the criticism will inspire hard work,” “he will be conceited if he believes he is the best,” and “if he believes he didn’t do well, he won’t be so careless next time”). Cohen’s kappa was 0.85.

A second pair of questions addressed children’s beliefs about the base rates of false positive and false negative feedback. The following example concerns false feedback that is positive.

Do you think a teacher might tell a student his work in writing is the best in the class if it’s not really true?

After first responding as to whether such a thing might happen, children who responded “yes” were asked, “does it happen hardly ever, sometimes, or most of the time?” This resulted in a *false feedback rating* with the following final response scale: *no* (scored as 0), *hardly ever* (scored as 1), *sometimes* (scored as 2), and *most of the time* (scored as 3).

## Results

Preliminary analyses showed no effects of participant gender, character gender, or subject area, so these variables were not included in subsequent analyses.

**Teacher credibility ratings.** Mean teacher credibility ratings are presented in Figure 2. A 2 (valence: positive, negative)  $\times$  2 (age group: younger, older)  $\times$  2 (location: U.S., China) ANOVA was conducted on mean credibility ratings of teacher feedback. The only significant effect was a main effect of valence,  $F(1, 124) = 37.42, p < .001$ , partial  $\eta^2 = 1.00$ , with greater acceptance of positive feedback ( $M = 2.26, SD = 0.81$ ) than negative feedback ( $M = 1.60, SD = 1.01$ ).

Follow-up analyses examining credibility ratings for praise and criticism separately indicate that there were no significant location differences in the willingness to accept praise, but that Chinese children were more likely to recommend accepting criticism ( $M = 1.78, SD = 0.98$ ) than were children from the U.S. ( $M = 1.39, SD = 1.00$ ).

An analysis of children’s open-ended explanations for their ratings of teacher credibility indicated that 39.0% of participants from the U.S. and 37.9% of those from China explicitly mentioned motivational factors on at least one of their two explanations. Within each country, most of the explanations made reference to the benefits of believing that one has high ability. However, explanations from Chinese participants were more likely to mention the costs of believing that one has high ability (34.2% of motivational explanations for Chinese participants vs. 13.8% of motivational explanations for participants from the U.S.).

**False feedback ratings.** If children believe that teachers are more likely to give false feedback that is negative than false feedback that is positive, it would seem logical for them to be more willing to accept negative feedback. However, the results from children’s false feedback ratings do not support this interpretation. Children from the U.S. tended to believe that teachers are more likely to give false feedback that is positive rather than negative,  $t(58) = 4.46, p < .001$  ( $M = 1.32, SD = 0.95$  for positive feedback;  $M = 0.66, SD = 0.73$  for negative feedback), and children from China viewed the possibility of teachers giving false positive or negative feedback as equally likely ( $M = 1.48$ , for positive feedback;  $M = 1.38, SD = 0.79$  for negative feedback).

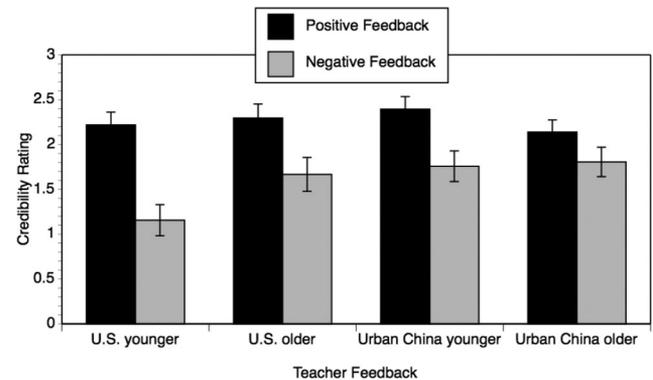


Figure 2. Mean credibility ratings of teacher feedback for each group in Study 2, for both positive and negative feedback. Error bars indicate standard errors.

## Discussion

Participants from both the U.S. and China were more skeptical of negative feedback than positive feedback, which replicated the central finding of Study 1. We also addressed the possibility that children would selectively reject negative feedback as a result of a possible assumption that most children tend to do very good work. We addressed this possibility by modifying the feedback to make it clear that a single student was being identified as either the best or the worst in the class. Because presumably there can be only one best or worst score within a particular class, assumptions about the base rate of successful or unsuccessful performance were no longer relevant to determining the accuracy of the feedback. That the selective skepticism effect replicated under these conditions suggests that it is not merely the product of a rational analysis of the available evidence.

Although participants in Study 2 showed evidence of selective skepticism, the effect was not as strong as in Study 1. It seems likely that the difference between the studies is due to the nature of the feedback. The feedback in Study 2 was more specific than in Study 1, so the participants in Study 2 may have felt less free to redefine success and failure in a way that allowed them to justify reaching a desirable conclusion.

The patterns of reasoning that were seen in Study 2 were more similar across age and cultural groups than those in Study 1. However, as in Study 1, children from the U.S. showed more skepticism about negative feedback than did children from China.

## General Discussion

Elementary school children were asked to judge the credibility of evaluative feedback concerning the academic performance of unfamiliar individuals. In each of the two studies, children in the U.S. and China showed a clear pattern of selective skepticism by rating negative feedback as less credible than positive feedback.

For children in the U.S., selective skepticism is consistent with the cultural emphasis on self-esteem (Miller et al., 2002) and self-promotion (Heine, 2001; Markus & Kitayama, 1991) that is prevalent in the West. The fact that selective skepticism was also seen in China suggests that selective skepticism may not be a culturally specific phenomenon. Although it is possible that Western influences in China have contributed to the selective skepticism that was evident in urban China, it is unlikely that such influences can fully account for the selective skepticism that was evident in rural China.

One possible explanation for the pattern of selective skepticism seen in Study 1 is that children were reasoning in a rational manner based on their assumptions about the quality of work done by most students. If children believe that most students do high-quality work, it might be logical to conclude that feedback judging a student's work to be "very good" is more likely to be true than feedback judging a student's work to be "not good at all." In Study 2, we explored this possibility by controlling for the relevant base rates. In particular, we presented evaluative feedback that described each student's performance as either the best in the class or the worst in the class. Even under these conditions, participants from both countries showed selective skepticism.

A related alternative explanation is that children tend to assume that evaluators are more likely to offer false negative feedback than false positive feedback. If this were indeed the case, selective

skepticism might constitute a rational response. However, results from Study 2 refute this possibility. It is particularly notable that children from the U.S. were inclined to believe positive feedback over negative feedback, even though they asserted that teachers are significantly more likely to give false positive feedback than false negative feedback.

We interpret the pattern of selective skepticism seen in the present study as evidence that when children assess the credibility of value-laden information, they have a tendency to consider not only objective information but also subjective factors such as the extent to which the information is desired. In particular, young children more readily believe positive evaluations about others over negative evaluations. This possibility is consistent with evidence of children's positivity biases (Boseovski, 2010; Boseovski & Lee, 2008; Heyman & Giles, 2004; Lockhart, Chang, & Story, 2002; Schuster et al., 1998) and parallels findings seen among adults, who often respond to undesirable information by engaging in deliberative efforts to discredit it (Ditto, Pizarro, & Tannenbaum, 2009; Uhlmann, Brescoll, & Pizarro, 2007).

Supporting our interpretation of our selective skepticism effect were children's explanations of their responses, in which references to motivation were often provided. With reference to praise, one child from the U.S. explained, "compared to everyone else, his was the best," but with reference to criticism declared, "it's just an opinion." Another child explained that, "when a teacher compliments work you should take pride in it" but that "if a teacher doesn't compliment your work, it's only one person so it doesn't mean much." These different ways of considering the evidence were also seen among the Chinese children. For example, one said, "you should believe it if it is good" when assessing praise but suggested "maybe the teacher made a mistake" when assessing criticism. Children's explanations also suggest that some children have a tendency to redefine success and failure so as to portray themselves and others favorably. For example, one child commented that negative feedback should be rejected because "maybe there is someone who writes worse than him," and another did so because "if you think your project is good, it is good."

Although we believe that our results support a utilitarian account of the selective skepticism effect, further research will be needed to confirm this possibility. Because children's open-ended explanations involved reflecting on their responses after the fact, there may be alternative accounts that could explain the present results. One plausible alternative account is a normative one, in which children accept positive feedback more readily than negative feedback because they tend to hear positive feedback more often. In fact, it would be surprising if normative factors did not play a role in shaping children's reasoning in this domain. It might also help to explain why the effects are especially strong in the U.S., where parents and teachers frequently seek to shield young children from exposure to overtly negative feedback (Stipek & Mac Iver, 1989; Twenge, 2006). However, it should be noted that there is no a priori reason to assume that the forms of feedback that are offered more frequently will necessarily be more accurate, given that some forms of feedback may be employed by teachers to affect student's emotions and motivation, rather than to make certain that students gain accurate information about their ability.

Another alternative account of the present results is that children might be more inclined to accept feedback when it is positive because it is easier to generate plausible explanations for positive

feedback than for negative feedback. As with the normative account, it is unclear whether this account would constitute a rational approach. In fact, if children tend to assume that people will accurately convey what they believe to be true, the ability to readily generate alternative explanations for evaluative feedback is likely to undermine this assumption.

Regardless of which interpretation of the selective skepticism effect ultimately prevails, the present results are likely to have important implications for children's developing conceptions of themselves and others. For example, previous research (e.g., Ehrlinger, 2012; Twenge, 2006) has suggested that people often give more positive feedback to others than is warranted, which in turn can lead people to develop overly positive views of themselves. Our findings suggest that children might develop inflated views of themselves not only because they hear positive feedback more often than negative feedback but also because they are less likely to question the validity of positive feedback than negative feedback.

Although our results demonstrate patterns of reasoning that extend beyond the U.S., they also point to significant cross-cultural differences. One such difference is that in both studies Chinese children were less skeptical about negative feedback from teachers than were children from the U.S. This difference may result from a greater focus within Chinese culture on the implications of negative feedback for future performance (Heyman et al., 2009), as well as a more general emphasis on learning as a means to perfect oneself morally and socially (Li, 2003, 2005). This emphasis may lead Chinese children to conclude that they should use negative feedback to better understand their own weaknesses and to identify what they must do to improve their future performance, rather than simply rejecting it because it is associated with unpleasant emotions.

Cognitive factors may eventually alter tendencies to engage in selective skepticism, with children's developing cognitive capacities allowing them to notice inconsistencies in their arguments and integrate relevant information more effectively (Schuster et al., 1998). For example, as children get older they may realize that if praise should be believed because "teachers always tell the truth," this reasoning should also be logically applied to cases that involve criticism. However, even when children have the required cognitive capacity to assess the credibility of information without regard to its desirability, they might not always do so. In the present study a small number of children made comments that suggested they might be struggling between what is logical and what is more desirable or intuitive. For example, after reporting that an instance of positive teacher feedback was credible, one older child paused before responding to the negative feedback item and said, "when it's positive, I would believe it, but it should be the same as when it's negative."

The extent to which the desirability of a positive or negative evaluation affects its perceived credibility may vary based on whether the individual making the assessment has a personal stake in the outcome. Children who are personally involved may be more likely to engage in wishful thinking in order to avoid reaching undesirable conclusions or experiencing negative emotions (Stipek, Roberts, & Sanborn, 1984). If the children in the present study had considered the outcome to be personally relevant there might have been stronger differences in reasoning across the samples. For example, children growing

up in China might be more willing than children in the U.S. to report that they would reject praise so as to appear modest (see Fu, Heyman, & Lee, 2011), a difference that might also be seen when comparing the responses of children growing up in rural China versus urban China (see Fu et al., 2010).

### Limitations and Future Directions

As noted previously, one limitation of the present study is that we were not able to definitively pin down the cause of the observed selective skepticism effect. There is also more work to be done to explain the effects of age and culture. In addition, it is not yet clear what role is played by specific social experiences, such as when a child observes a teacher repeatedly presenting evaluative feedback that is clearly inaccurate (Birch et al., 2008; Corriveau & Harris, 2000; Jaswal & Neely, 2006; Koenig & Harris, 2005) or acting in a manner that suggests he or she is biased (Mills & Keil, 2005, 2008; Spears Brown & Bigler, 2005). Although some studies have sought to characterize children's experiences with evaluative feedback (e.g., Ng et al., 2007; Stigler & Perry, 1990), more work will be needed to gain a comprehensive picture of these experiences and how they might relate to reasoning about evaluative feedback. Further research will also be needed to explain the cross-cultural differences that were seen in relation to cultural values. For example, the cultural emphasis on respecting elders in China (Domino, Affonso, & Slobin, 1987) might make Chinese children more likely to accept negative messages from teachers.

It will also be important for future researchers to address individual differences in children's perceptions of credibility. As groups, children from the U.S. and from China were more skeptical of criticism than of praise, but a small number within each country showed the reverse pattern. One child from the U.S. argued that teachers sometimes offer praise to make students feel better but that when they offer criticism it means that the work is not "up to standard." This raises the possibility that there may be systematically different ways of evaluating evidence in this domain that have meaningful psychological consequences, similar to what has been found with reference to children's tendency to overestimate versus underestimate their own abilities (Eshel & Kurman, 1991).

In the present study, we used measures that involved holding scenarios constant in every way except for the variables of interest. This approach may have called attention to the contrast between positive and negative evaluative feedback in a way that is unlikely to occur in children's social experiences. In future research, it will be important to develop measures to better assess how children apply these capacities within school and home contexts.

### Conclusion

Children from the U.S. and China were less willing to accept evaluative feedback when it was negative than when it was positive, which we described as the selective skepticism effect. Although the degree of selective skepticism varied as a function of age and population, it was seen when children assessed feedback from teachers and from peers. The findings suggest that children's decisions about what information to believe may be influenced by the anticipated psychological consequences of the information itself.

## References

- Aloise-Young, P. A. (1993). The development of self-presentation: Self-promotion in 6- to 10-year-old children. *Social Cognition, 11*, 201–222. doi:10.1521/soco.1993.11.2.201
- Altermatt, E. R., Pomerantz, E. M., Ruble, D. N., Frey, K. S., & Greulich, F. K. (2002). Predicting changes in children's self-perceptions of academic competence: A naturalistic examination of evaluative discourse among classmates. *Developmental Psychology, 38*, 903–917. doi:10.1037/0012-1649.38.6.903
- Banerjee, R. (2000). The development of an understanding of modesty. *British Journal of Developmental Psychology, 18*, 499–517. doi:10.1348/026151000165823
- Banerjee, R., & Yuill, N. (1999). Children's explanations for self-presentational behaviour. *European Journal of Social Psychology, 29*, 105–111. doi:10.1002/(SICI)1099-0992(199902)29:1<105::AID-EJSP910>3.0.CO;2-K
- Benne, M., & Yeeles, C. (1990). Children's understanding of the self-presentational strategies of ingratiation and self-promotion. *European Journal of Social Psychology, 20*, 455–461. doi:10.1002/ejsp.2420200507
- Birch, S. A., Vauthier, S. A., & Bloom, P. (2008). Three- and four-year-olds spontaneously use others' past performance to guide their learning. *Cognition, 107*, 1018–1034. doi:10.1016/j.cognition.2007.12.008
- 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. doi:10.1111/j.1750-8606.2010.00149.x
- 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
- Boseovski, J. J., & Lee, K. (2008). Seeing the world through rose-colored glasses? Neglect of consensus information in young children's personality judgments. *Social Development, 17*, 399–416. doi:10.1111/j.1467-9507.2007.00431.x
- Chen, X., Cen, G., Li, D., & He, Y. (2005). Social functioning and adjustment in Chinese children: The imprint of historical time. *Child Development, 76*, 182–195. doi:10.1111/j.1467-8624.2005.00838.x
- Chen, X., Wang, L., & Wang, Z. (2009). Shyness-sensitivity and social, school, and psychological adjustment in rural migrant and urban children in China. *Child Development, 80*, 1499–1513. doi:10.1111/j.1467-8624.2009.01347.x
- Corriveau, K., & Harris, P. L. (2009). Preschoolers continue to trust a more accurate informant 1 week after exposure to accuracy information. *Developmental Science, 12*, 188–193. doi:10.1111/j.1467-7687.2008.00763.x
- Ditto, P. H., Pizarro, D. A., & Tannenbaum, D. (2009). Motivated moral reasoning. In B. H. Ross (Series Ed.) & D. M. Bartels, C. W. Bauman, L. J. Skitka, & D. L. Medin (Vol. Eds.), *Psychology of learning and motivation: Vol. 50. Moral judgment and decision making* (pp. 307–338). San Diego, CA: Academic Press.
- Domino, G., Affonso, D., & Slobin, M. (1987). Community psychology in the People's Republic of China. *Psychologia: An International Journal of Psychology in the Orient, 30*, 1–11.
- Dweck, C. S. (1999). *Self-theories: Their role in motivation, personality, and development*. Philadelphia, PA: Psychology Press.
- Ehrlinger, J. (2012, August). Polite but not honest: How an absence of negative social feedback contributes to overconfidence. In J. Ehrlinger (Chair), *Sources of accuracy and error in social judgment*. Symposium conducted at the meeting of the American Psychological Association, Orlando, FL.
- Eshel, Y., & Kurman, J. (1991). Academic self-concept, accuracy of perceived ability and academic attainment. *British Journal of Educational Psychology, 61*, 187–196. doi:10.1111/j.2044-8279.1991.tb00974.x
- Fu, G., Brunet, M. K., Lv, Y., Ding, X., Heyman, G. D., Cameron, C. A., & Lee, K. (2010). Chinese children's moral evaluation of lies and truths—roles of context and parental individualism–collectivism tendencies. *Infant and Child Development, 19*, 498–515.
- Fu, G., Heyman, G. D., & Lee, K. (2011). Reasoning about modesty among adolescents and adults in China and the U.S. *Journal of Adolescence, 34*, 599–608. doi:10.1016/j.adolescence.2010.10.003
- Gelman, S. A. (2009). Learning from others: Children's construction of concepts. *Annual Review of Psychology, 60*, 115–140. doi:10.1146/annurev.psych.59.103006.093659
- Harris, P. L. (2007). Trust. *Developmental Science, 10*, 135–138. doi:10.1111/j.1467-7687.2007.00575.x
- Heine, S. J. (2001). Self as cultural product: An examination of East Asian and North American selves. *Journal of Personality, 69*, 881–905. doi:10.1111/1467-6494.696168
- Helwig, C. C., Yang, S., Tan, D., Liu, C., & Shao, T. (2011). Urban and rural Chinese adolescents' judgments and reasoning about personal and group jurisdiction. *Child Development, 82*, 701–716. doi:10.1111/j.1467-8624.2010.01531.x
- Hess, R., Chang, C., & McDevitt, T. (1987). Cultural variations in family beliefs about children's performance in mathematics: Comparisons among People's Republic of China, Chinese-American, and Caucasian-American families. *Journal of Educational Psychology, 79*, 179–188. doi:10.1037/0022-0663.79.2.179
- Heyman, G. D. (2008). Children's critical thinking when learning from others. *Current Directions in Psychological Science, 17*, 344–347. doi:10.1111/j.1467-8721.2008.00603.x
- Heyman, G. D., Fu, G., & Lee, K. (2007). Evaluating claims people make about themselves: The development of skepticism. *Child Development, 78*, 367–375. doi:10.1111/j.1467-8624.2007.01003.x
- Heyman, G. D., Fu, G., Sweet, M. A., & Lee, K. (2009). Children's reasoning about evaluative feedback. *British Journal of Developmental Psychology, 27*, 875–890. doi:10.1348/026151008X390870
- Heyman, G. D., & Giles, J. W. (2004). Valence effects in reasoning about evaluative traits. *Merrill-Palmer Quarterly, 50*, 86–109. doi:10.1353/mpq.2004.0004
- Heyman, G. D., & Legare, C. H. (2005). Children's evaluation of sources of information about traits. *Developmental Psychology, 41*, 636–647. doi:10.1037/0012-1649.41.4.636
- Jaswal, V. K., & Neely, L. A. (2006). Adults don't always know best: Preschoolers use past reliability over age when learning new words. *Psychological Science, 17*, 757–758. doi:10.1111/j.1467-9280.2006.01778.x
- Kelley, H. H. (1973). The processes of causal attribution. *American Psychologist, 28*, 107–128. doi:10.1037/h0034225
- Koenig, M. A., & Harris, P. L. (2005). Preschoolers mistrust ignorant and inaccurate speakers. *Child Development, 76*, 1261–1277. doi:10.1111/j.1467-8624.2005.00849.x
- Lee, K., & Cameron, C. A. (2000). Extracting truthful information from lies: Emergence of the expression–representation distinction. *Merrill-Palmer Quarterly, 46*, 1–20.
- Li, J. (2003). The core of Confucian learning. *American Psychologist, 58*, 146–147. doi:10.1037/0003-066X.58.2.146
- Li, J. (2005). Mind or virtue: Western and Chinese beliefs about learning. *Current Directions in Psychological Science, 14*, 190–194. doi:10.1111/j.0963-7214.2005.00362.x
- Lockhart, K. L., Chang, B., & Story, T. (2002). Young children's beliefs about the stability of traits: Protective optimism? *Child Development, 73*, 1408–1430. doi:10.1111/1467-8624.00480
- Markus, H. R., & Kitayama, S. (1991). Culture and the self: Implications for cognition, emotion, and motivation. *Psychological Review, 98*, 224–253. doi:10.1037/0033-295X.98.2.224
- Miller, P. J., Wang, S., Sandel, T., & Cho, G. E. (2002). Self-esteem as folk theory: A comparison of European American and Taiwanese mothers' beliefs. *Parenting: Science and Practice, 2*, 209–239. doi:10.1207/S15327922PAR0203\_02

- Mills, C. M., & Grant, M. G. (2009). Biased decision-making: Developing an understanding of how positive and negative relationships may skew judgments. *Developmental Science, 12*, 784–797. doi:10.1111/j.1467-7687.2009.00836.x
- Mills, C. M., & Keil, F. C. (2005). The development of cynicism. *Psychological Science, 16*, 385–390. doi:10.1111/j.0956-7976.2005.01545.x
- Mills, C. M., & Keil, F. C. (2008). Children's developing notions of (im)partiality. *Cognition, 107*, 528–551. doi:10.1016/j.cognition.2007.11.003
- Moses, L. J., & Baldwin, D. A. (2005). What can the study of cognitive development reveal about children's ability to appreciate and cope with advertising? *Journal of Public Policy & Marketing, 24*, 186–201. doi:10.1509/jppm.2005.24.2.186
- Ng, F. F., Pomerantz, E. M., & Lam, S. (2007). Chinese and European American parents' responses to children's success and failure: Implications for children's responses. *Developmental Psychology, 43*, 1239–1255. doi:10.1037/0012-1649.43.5.1239
- Parsons, J. E., Kaczala, C. M., & Meece, J. L. (1982). Socialization of achievement attitudes and beliefs: Classroom influences. *Child Development, 53*, 322–339. doi:10.2307/1128974
- Robinson, E. J., Champion, H., & Mitchell, P. (1999). Children's ability to infer utterance veracity from speaker informedness. *Developmental Psychology, 35*, 535–546. doi:10.1037/0012-1649.35.2.535
- Sabbagh, M. A., & Baldwin, D. A. (2001). Learning words from knowledgeable versus ignorant speakers: Links between preschoolers' theory of mind and semantic development. *Child Development, 72*, 1054–1070. doi:10.1111/1467-8624.00334
- Schuster, B., Ruble, D., & Weinert, F. (1998). Causal inferences and the positivity bias in children: The role of the covariation principle. *Child Development, 69*, 1577–1596.
- Spears Brown, C., & Bigler, R. S. (2005). Children's perception of discrimination: A developmental model. *Child Development, 76*, 533–553. doi:10.1111/j.1467-8624.2005.00862.x
- Stigler, J. W., & Perry, M. (1990). Mathematics learning in Japanese, Chinese, and American classrooms. In J. W. Stigler, R. A. Shweder, & G. Herdt (Eds.), *Cultural psychology: Essays on comparative human development* (pp. 328–353). Cambridge, England: Cambridge University Press.
- Stipek, D. J., & Mac Iver, D. J. (1989). Developmental change in children's assessment of intellectual competence. *Child Development, 60*, 521–538. doi:10.2307/1130719
- Stipek, D. J., Roberts, T. A., & Sanborn, M. E. (1984). Preschool-age children's performance expectations for themselves and another child as a function of the incentive value of success and the salience of past performance. *Child Development, 55*, 1983–1989. doi:10.2307/1129773
- Twenge, J. M. (2006). *Generation me*. New York, NY: Free Press.
- Uhlmann, E. L., Brescoll, V. L., & Pizarro, D. A. (2007). The motivated use and neglect of base rates. *Behavioral and Brain Sciences, 30*, 284–285. doi:10.1017/S0140525X07001938
- Vanderbilt, K. E., Liu, D., & Heyman, G. D. (2011). The development of distrust. *Child Development, 82*, 1372–1380. doi:10.1111/j.1467-8624.2011.01629.x
- Wang, F., Liu, S., & Yinf, F. (2009). *Why don't we praise?* Hangzhou, China: Zhejiang Educational Press.
- Watling, D., & Banerjee, R. (2007). Children's understanding of modesty in front of peer and adult audiences. *Infant and Child Development, 16*, 227–236. doi:10.1002/icd.450
- Woolley, J. D., Boerger, E. A., & Markman, A. (2004). A visit from the Candy Witch: Children's belief in a novel fantastical entity. *Developmental Science, 7*, 456–468. doi:10.1111/j.1467-7687.2004.00366.x
- Zhang, W., & Fuligni, A. J. (2006). Authority, autonomy, and family relationships among adolescents in urban and rural China. *Journal of Research on Adolescence, 16*, 527–537. doi:10.1111/j.1532-7795.2006.00506.x

## Appendix A

### Scenarios

#### Example Scenarios From Study 1

Student and teacher names were selected from a pool and assigned randomly. The gender of the characters was counterbalanced. Participants heard versions of all eight of the following scenarios.

*positive feedback, teacher, single student*

Ryan wrote an essay about a class field trip. His teacher, Mr. Turner, read Ryan's essay before reading anyone else's and told Ryan that his essay is very good. Should Ryan believe that his essay is very good?

*positive feedback, teacher, all individuals in class*

Luke wrote an essay about a class field trip. His teacher, Mr. Peterson, first read all of the other kids' essays and told all of them that their essays are very good. Then, Mr. Peterson also told Luke that his essay is very good. Should Luke believe that his essay is very good?

*negative feedback, teacher, single student*

John wrote an essay about a class field trip. His teacher, Mr. Watkins, read John's essay before reading anyone else's and told John that his essay is not good at all. Should John believe that his essay is not good at all?

(Appendices continue)

*negative feedback, teacher, all individuals in class*

Christopher wrote an essay about a class field trip. His teacher, Mr. Jones, first read all of the other kids' essays and told all of them that their essays are not good at all. Then, Mr. Jones also told Christopher that his essay is not good at all. Should Christopher believe that his essay is not good at all?

*positive feedback, classmate, single student*

Thomas wrote an essay about a class field trip. His classmate, Michael, read Thomas's essay before reading anyone else's and told Thomas that his essay is very good. Should Thomas believe that his essay is very good?

*positive feedback, classmate, all individuals in class*

Oliver wrote an essay about a class field trip. His classmate, Nathan, first read all of the other kids' essays and told all of them that their essays are very good. Then, Nathan also told Oliver that his essay is very good. Should Oliver believe that his essay is very good?

*negative feedback, classmate, single student*

Anthony wrote an essay about a class field trip. His classmate, Eric, read Anthony's essay before reading anyone else's and told Anthony that his essay is not good at all. Should Anthony believe that his essay is not good at all?

*negative feedback, classmate, all individuals in class*

Joshua wrote an essay about a class field trip. His classmate, Daniel, first read all of the other kids' essays and told all of them that their essays are not good at all. Then, Daniel also told Joshua that his essay is not good at all. Should Joshua believe that his essay is not good at all?

## Example Scenarios From Study 2

Student and teacher names were selected from a pool and assigned randomly. The gender of the characters was counterbalanced. Participants heard either the two scenarios in Set A or the two scenarios in Set B.

*Set A, praise*

David wrote an essay for his homework. His teacher, Mr. Peterson, told David that his essay was the best in the class. Should David believe Mr. Peterson that his essay was the best in the class?

*Set A, criticism*

Christopher wrote an essay for his homework. His teacher, Mr. Turner, told Christopher that his essay was the worst in the class. Should Christopher believe Mr. Peterson that his essay was the worst in the class?

*Set B, praise*

Jacob wrote a science project for his homework. His teacher, Mr. Peterson, told Jacob that his science project was the best in the class. Should Jacob believe Mr. Peterson that his science project was the best in the class?

*Set B, criticism*

Nicholas wrote a science project for his homework. His teacher, Mr. Turner, told Nicholas that his science project was the worst in the class. Should Nicholas believe Mr. Peterson that his science project was the worst in the class?

## Appendix B

### Examples of Open-Ended Explanations

Examples of explanations that contain explicit references to motivation or desirability factors.

She shouldn't believe what people say when they criticize. She should only believe praise.

It's good to feel that you're the best.

He should believe it if it makes him feel good.

Never think that your work is horrible.

She should believe in herself.

If he believes it, he will lose confidence.

You shouldn't believe you're too good or you could be tricked and not practice more.

You shouldn't think too highly of yourself.

Examples of explanations without explicit reference to motivation or desirability factors.

You should believe what teachers say.

It's the teacher's opinion and the teacher is the boss of you.

Maybe the teacher thought that he was the best.

Teachers make mistakes sometimes.

He is the most clever student in the class.

Maybe he is a really good writer.

She wasn't up to the standard.

Someone could have done worse.

Maybe he offended the teacher in his essay, so the teacher doesn't want to give him a high grade.

I had a personal experience where my math teacher praised me, but I wasn't all that good.

Received May 19, 2011

Revision received August 29, 2012

Accepted October 16, 2012 ■