ELSEVIER

Children's developing notions of (im)partiality

Candice M. Mills [a],*, Frank C. Keil [b]

[a] *The University of Texas at Dallas, P.O. Box 830688, GR41, Richardson, TX 75083, USA*
[b] *Department of Psychology, Yale University, United States*

Abstract

This research examines the development of children's understanding that people's judgments may be skewed by relationships, and that situational factors may make it difficult to be impartial. One hundred and seventy-one adults and children between kindergarten and eighth grade heard stories about judges in contests with objective or subjective criteria for winning. In Experiment 1, by fourth grade, children rated a judge with no personal connection (the "neutral judge") as being more likely to be objective than a judge with a personal connection (the "connected judge"). Younger children showed the opposite pattern. Experiment 2 replicated this finding for judges, and also found that children across development have similar ideas regarding the characteristics for being a good judge. Not until eighth grade, however, did children indicate that a connected judge was more problematic in subjective situations than in objective ones.
Published by Elsevier B.V.

Keywords: Cognitive development; Social cognition; Judgment; Epistemology; Critical thinking; Skepticism; Children

1. Introduction

Standing in front of the United States Senate as a nominee for the Supreme Court, Samuel Alito was recently inundated with questions regarding his political

---
* Corresponding author. Tel.: +1 972 883 4475; fax: +1 972 883 2491.
  E-mail address: candice.mills@utdallas.edu (C.M. Mills).

background and his agenda for sitting on the highest court in the country. Time and time again, senators asked Alito about his qualifications in light of his politics, clearly concerned that Alito's conservative leanings might color his judgments regarding the future of several key laws.

In everyday life, people often show concern about others' qualifications to make judgments about information. In court cases, people with personal connections to the criminals are not allowed to serve on the jury. In international athletic contests, great effort is taken to ensure that judges do not give significantly higher scores to contestants from their own countries, often by preventing them from judging their compatriots. In academic settings, researchers are prevented from evaluating grants and research articles from their colleagues. People assume that judges in these situations may be poor evaluators of information, implicitly biased by their preexisting beliefs or preferences, or that they may be unfair in their judgments, even to the point of deliberate deception. This article examines how intuitions regarding partiality as well as other characteristics of being a good judge emerge in development.

Why do we care about the development of an understanding of partiality? Children and adults are bombarded with information from many different sources, and not all of these sources have pure intentions or provide accurate messages. Being able to predict that someone may make an inaccurate or skewed claim may help dampen the negative effects of such claims. In research with adults, for instance, the timing of discovering that testimony may be inaccurate influences how they remember the information. If adults receive some sort of discounting cue before being exposed to a persuasive message, they tend to process the information more cautiously. Alternatively, if they receive a discounting cue after a message, they often forget that the information came from an inaccurate source and just focus on the information itself (Petty & Cacioppo, 1986). In another line of research, children who monitor the source of acquired information are less suggestible than they would be otherwise (Thierry & Spence, 2002; Thierry, Spence, & Memon, 2001). Thus, by knowing in advance that a source may be partial and therefore less trustworthy, children may be better at coping with potentially misleading information. Most adults understand the above examples of concern regarding the possibility of partial judgments, but it is an open question how this understanding develops.

In order to be able to determine if someone might make partial or unfair judgments, it is first important to have some concept of what it means to be unfair. Preschoolers seem to define fairness in terms of the equal distribution of resources (e.g., each child in a classroom should get the same number of gold stars for completing a task, regardless of how well the task was completed). By kindergarten, however, children can see fairness in terms of merit – that rewards are often commensurate with quality (Damon, 1988; Lerner, 1974; Nelson & Dweck, 1977). Although there are cultural and situational differences in what is deemed the appropriate way to distribute resources (e.g., Rawls, 1971), understanding that some people are more worthy of a reward than others is essential in being able to evaluate whether or not a judgment has been fair.

After understanding fairness, one must know what characteristics qualify someone to be an excellent judge, and what characteristics may make someone unable

or unwilling to offer a fair and objective judgment. Research on how children determine fair procedures for offering rewards focuses on the needs of the contestants in contests, or the requirements for winning the contests (e.g., skill versus luck), not on the qualifications of the judges themselves (e.g., Sigelman & Waitzman, 1991; Thorkildsen & White-McNulty, 2002). However, related research examining how children evaluate information suggests that young children recognize the importance of informed testimony and judgments: They distinguish knowledgeable speakers from ignorant ones when learning new words (Jaswal & Neely, 2006; Koenig, Clement, & Harris, 2004; Sabbagh & Baldwin, 2001) and when dealing with misinformation (Welch-Ross, 1999). Four-year-olds understand that sometimes people do not have enough knowledge to provide accurate statements (Sodian, 1988; Wimmer & Perner, 1983), and that some people are more knowledgeable than others for a given topic (Lutz & Keil, 2002). Preschoolers discount information presented by a discredited source, such as an adult who was described as being ''silly'' (Lampinen & Smith, 1995), or someone that has made a series of mistakes in the past (Pasquini, Corriveau, Koenig, & Harris, 2007), and kindergartners discount implausible statements about transgressions that clearly contradict with reality (e.g., a ghost jumped out of a book and broke a glass; Lee, Cameron, & Doucette, 2002).

Thus, young children can recognize that someone can be a poor judge, and may be unable to provide a correct assessment. But what about in cases where it is less certain, but highly likely, that someone may show poor judgment? In cases where judges have a personal connection to the judgment at hand (for example, by personal relationship or country affiliation), adults believe that the likelihood of an objective judgment is greatly reduced; the risk of the judges making a biased decision, or even lying, seems too high. For example, adults assume that people's attitudes (e.g., on abortion rights) are strongly influenced, and perhaps even skewed, by group membership (e.g., gender; Miller & Ratner, 1998). People expect judgments regarding the allocation of responsibility to be motivationally biased (Kruger & Gilovich, 1999). People also expect others to act deceptively in some situations, such as in politics (McGraw, Lodge, & Jones, 2002), or when ulterior motives are clearly present (Fein, 1996).

While it is unknown when children understand that someone might be biased or might lie for someone else (e.g., a friend, a member of their social group) to receive a reward, there is evidence that elementary school aged children recognize that personal motivations might corrupt beliefs and statements. By fourth grade, children discount self-report as a source of information for learning about evaluative traits in other people, concerned that such reports might be skewed by a desire to be seen in a positive way (Heyman, Fu, & Lee, 2007; Heyman & Legare, 2005). By second grade, children discount others' self-interested statements regarding the outcomes of contests, recognizing that their statements may be skewed by their motivations to win a prize (Mills & Keil, 2005). Even younger children show some grasp of manipulative intent, understanding that there are times when people deceive others to get what they want (Bussey, 1992; Robinson, Mitchell, & Nye, 1995). This research suggests in middle childhood, children understand that self-interests can influence what people say, but it does not reveal how much children understand regarding favoritism or partiality.

Understanding that relationships can influence judgments requires first recognizing that judgments can vary between people. Pillow (1991) found that kindergartners and second graders use information about a character's preexisting biases (either positive or negative) to determine how someone will construe an ambiguous action, given several options. When no information is given about someone's preexisting bias, it is not until second grade when children accept that two different interpretations regarding the same ambiguous stimuli are reasonable (Carpendale & Chandler, 1996). Also around this age, children recognize that a person's personality traits will influence their future thoughts and behavior (Gnepp & Chilamkurti, 1988; Heyman & Gelman, 1999; Yuill & Pearson, 1998). Thus, at some point during the early elementary school years, children seem to be able to use someone's past behavior or preexisting beliefs to predict some future behaviors.

Contextual factors can also influence the degree to which personal connections produce partial judgments. In situations with clear criteria for judgments, there is less concern about personal interests unintentionally skewing outcomes. For a running race, for example, the person who crosses the finish line first should win, and unless the race is extremely close, it is difficult for a judge to be partial. For a beauty contest, on the other hand, where judgments of what qualifies as "most beautiful" are debatable, adults are more concerned if a contestant's closest friend, parent, or teacher is one of the judges. Even early in childhood, children recognize that there are situations in which people may have different tastes and preferences (Nichols & Folds-Bennett, 2003). Other research has found that even by kindergarten, children recognize that judges are more likely to disagree in contests with more subjective criteria (e.g., baking contest, art contest) than those with more objective criteria (e.g., running race, spelling bee) (Mills, 2006, 2007). While previous research may make it clear that children can distinguish some matters of taste from matters of fact by kindergarten or second grade (Carpendale & Chandler, 1996; Flavell, Flavell, Green, & Moses, 1990; Kuhn, Cheney, & Weinstock, 2000; Kuhn & Weinstock, 2002), it is unknown if children recognize that judgments are more likely to be skewed in subjective contexts compared to objective ones.

Therefore, we ask three central questions here. First, do children recognize that self-interests in the form of personal relationships might influence how people judge the outcomes of contests, either intentionally or unconsciously? Second, when do children come to understand that some situations require impartial judges more than others? In Experiment 1, participants heard stories involving contests with objective criteria (e.g., running race, pie eating contest) and contests with more subjective criteria (e.g., beauty contest, talent contest). For each story, participants evaluated two potential judges, one with a personal connection to the character in the contest (e.g., the character's parent, teacher, or best friend), and one who did not know anyone in the contest personally. Given that previous research has found that by second to fourth grade, children understand that preexisting biases might influence what someone says or believes (e.g., Heyman & Legare, 2005; Mills & Keil, 2005; Pillow, 1991), we predicted that by those grades, children would recognize the conflict in using judges with personal connections. However, recognizing that judges with connections are more likely to make inaccurate judgments in subjective contests compared

to objective ones may be a more difficult problem, so we predicted that only older children and adults would make this distinction.

We examined these questions using a slightly different format in Experiment 2, along with a third, more general question: what criteria do children think are important for being a good judge, and how does that change across development?

2. Experiment 1

2.1. Methods

2.1.1. Participants

Eighteen kindergartners (mean age = 5 years 11 months; range = 5 years 2 months to 6 years 10 months), 18 second graders (mean age = 8 years 0 months; range = 7 years 6 months to 8 years 6 months), and 18 fourth graders (mean age = 9 years 10 months; range = 9 years 6 months to 10 years 6 months). The sample was gender balanced and reflected the distribution of ethnic groups in the community: approximately 82% Caucasian, 12% African American, and 6% of other races. Children were recruited from the greater New Haven area and were tested in a quiet room; each session took about 15 min.

Additionally, 18 sixth graders (age 11 or 12), 18 eighth graders (age 13 or 14), and 18 adults also participated in the same study on paper. The older children were recruited from the same school district as the elementary school children, and the adults were undergraduate students in the New Haven area. This sample was also gender balanced.

2.1.2. Design

Participants heard two different kinds of stories: six stories involved objective events (running race, tower building contest, math contest, pie eating contest, long jump contest, spelling bee), while six involved subjective events (beauty contest, dance contest, talent show, story-writing contest, baking contest, art contest). For each story, participants were told about two potential judges: one with a personal connection to the character in the contest (e.g., the character's parent, teacher, or best friend), and one who did not know anyone in the contest personally (but was still a parent, teacher, or child of the same age group). There were three different possible targets (a parent, a teacher, or a best friend). Therefore, this was a 2 (judge connection) × 2 (contest type) × 3 (target) design.

All stories were about four sentences long. The stories were placed into four packets, pseudo-randomly ordered. See Appendix A for sample stories.

2.1.3. Procedure

Prior to testing, participants were asked if they had ever seen or been in a contest before. The experimenter told the child that, for contests, it is important to have a good judge to determine the winner. Children were told that in some situations, it is easier to judge fairly than in other situations, and to be a fair judge, you should

make your decision about the winner of the contest based *only* on that contest, and nothing else. A fair judge should give the prize to the person who did the best job at the contest they were in.

Children were trained to use a scale of 1–5 stars to rate how likely it was for each judge to pick the person who really met the criteria to win the race, with 1 meaning the judge would do a poor job, and 5 meaning the judge would do a good job. For training, they were told about three judges that represented different points on the scale. For an English poetry contest, one judge read and spoke only Chinese; children were then asked, "How fair would that judge be? How good would he be at picking out who had the best poem?" This judge represented the 1-star rating. For a diving contest, children were told about two judges: one judge knew a lot about diving (5-stars), while another judge knew about diving but brought his cell phone and a book with him to read in case he got bored (2 or 3 stars). All children were able to successfully complete the training, demonstrating an understanding of the different levels of ratings.

After training, children were told that they would hear about some contests, and that they would be asked to think about how good different judges would be. The experimenter then read the stories, periodically asking fact-check questions about the contest. For each of the 12 contests, a drawing was placed on the table to maintain attention. The drawings referred to the topic of the story, but importantly, these drawings did not provide information about the outcome of the event. Following each story, participants heard about two judges: one with a personal connection (the connected judge) and one without a connection (the neutral judge). Two drawings of stick figures were placed on the table to represent the targets: a teacher stood in front of a blackboard, a parent wore either a skirt or a tie, and a peer wore a tee shirt. The stick figures wore different color shirts so the children could tell them apart, but they were otherwise the same. Children were asked to use the scale to rate the judges based on the questions, "How fair would that judge be? How good would he be at picking out who really [met the criteria for being the winner of that particular contest (e.g., crossed the finish line first in the running race; was the most beautiful girl in the beauty contest)]?" The judges were presented in random order to the children. Half the judges were male and half were female.

The pencil-and-paper version for sixth graders, eighth graders, and adults included a set of written instructions with the accompanying examples and pictures from the experimenter script. The procedure was otherwise the same.

2.2. Results

For each contest type and target, we calculated the average rating for proposed judge fairness across judges and stories for participants in each grade as well as adults. See Table 1 for participants' mean ratings for the different contest types and judge connection. There were no differences in mean responses between the six stories with judges in contests with objective criteria ("objective contests") or between the six stories with judges in contests with more subjective criteria ("subjective contests"), so the data from the categories of objective and subjective contests

Table 1
Means and standard errors for ratings of fairness for connected and neutral judges in objective and subjective contests

Grade	Objective contests		Subjective contests		Collapsed across contest type	
	Connected judge	Neutral judge	Connected judge	Neutral judge	Connected judge	Neutral judge
Kindergarten	4.15 (0.24)	2.57 (0.31)	4.07 (0.19)	2.68 (0.23)	4.11 (0.20)	2.62 (0.28)
Second	3.22 (0.20)	3.60 (0.31)	3.27 (0.25)	3.59 (0.31)	3.25 (0.23)	3.60 (0.30)
Fourth	2.69 (0.20)	4.36 (0.16)	2.49 (0.18)	4.32 (0.20)	2.59 (0.18)	4.34 (0.18)
Sixth	2.18 (0.21)	4.72 (0.09)	2.05 (0.18)	4.73 (0.08)	2.11 (0.19)	4.73 (0.08)
Eighth	2.56 (0.23)	4.84 (0.09)	2.19 (0.16)	4.82 (0.09)	2.37 (0.18)	4.82 (0.09)
Adults	3.48 (0.25)	4.50 (0.07)	2.26 (0.17)	4.52 (0.12)	2.92 (0.10)	4.51 (0.06)
Total	3.04 (0.09)	4.10 (0.08)	2.75 (0.08)	4.06 (0.08)	2.89 (0.08)	4.08 (0.08)

Note. Mean ratings are based on a 5-point scale; standard errors are in parentheses.

will be reported here as averages for each of the two categories as a whole. There was also no effect of sex for any of the analyses, so it was excluded from further analyses.

A Repeated Measures ANOVA with judge rating, contest type, and target as within-subject factors and grade as a between-subjects variable was conducted. Overall, there was no significant difference in the between-subjects factor, mean ratings across grades, $F(5, 102) = 1.131$, $p = .349$. Descriptively, kindergartners had the lowest mean rating for judges ($M = 3.363$), and as children got older, their mean ratings were higher, with adults having the highest mean rating ($M = 3.734$).

Overall, participants rated judges in objective contests as being more fair than judges in subjective contests, $F(1, 102) = 26.194$, $p < .001$, $\eta^2 = .204$. There was also a contest type by grade interaction, $F(5, 102) = 10.542$, $p < .001$, $\eta^2 = .341$. Kindergartners and second graders did not give different ratings to judges in the two types of contests, but older children and adults gave higher ratings to judges in objective contests than subjective ones, regardless of judge connection.

Participants distinguished between the targets, rating teachers the highest, followed by parents and then friends, $F(2, 204) = 24.429$, $p < .001$, $\eta^2 = .193$. There was also a target by grade interaction, in that older children and adults were more likely to rate teachers higher than parents and peers, while kindergartners and second graders rated teachers and parents similarly $F(10, 204) = 2.770$, $p < .01$, $\eta^2 = .120$ (see Fig. 1).

Overall, participants rated the judges with no connection to the character in the story ("neutral judges") as being more fair than the judges with a connection ("connected judges"), $F(1, 102) = 83.796$, $p < .001$, $\eta^2 = .451$. To examine when children recognize that self-interests in the form of personal relationships might influence how people judge the outcomes of contests, we looked at the interaction with judge rating and grade (shown in Fig. 2). There was indeed an interaction, indicating different patterns of responses across development, $F(5, 102) = 23.459$, $p < .001$, $\eta^2 = .535$. Kindergartners rated connected judges higher than neutral ones, second graders rated them about equally, and older children and adults rated neutral judges
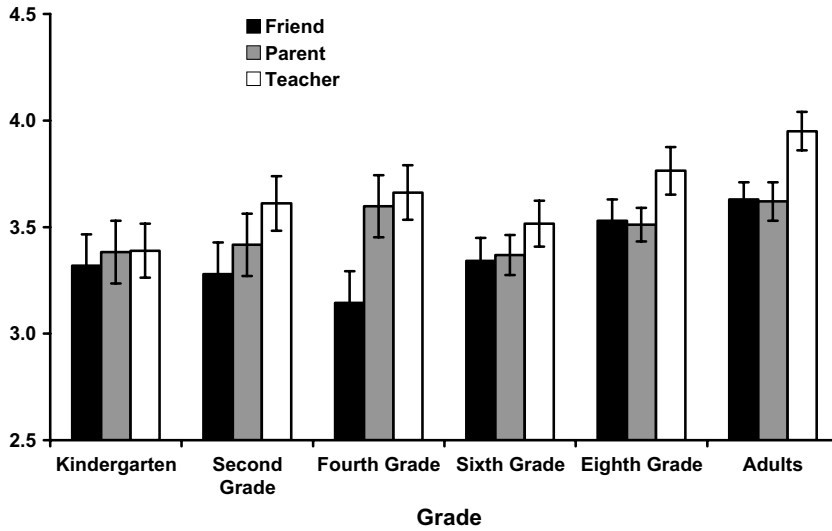
Fig. 1. Ratings for target of friend, parent, and teacher, collapsed across contest type and judge.
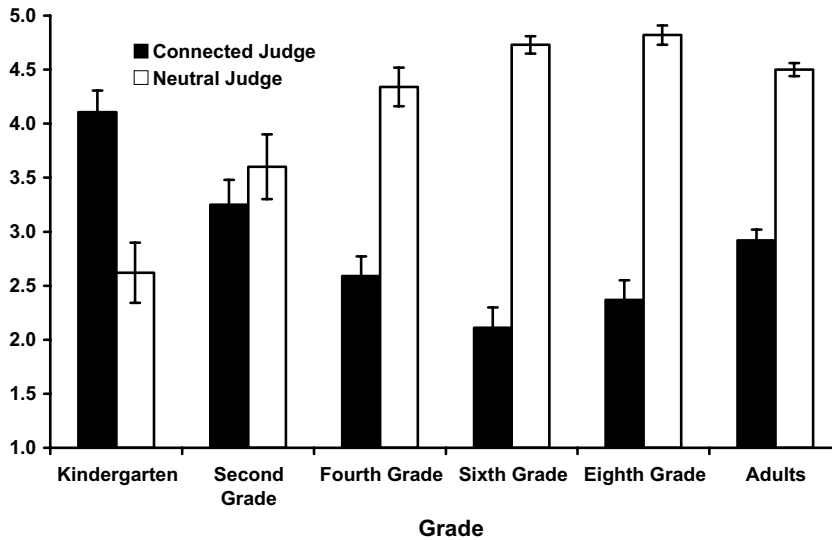

Fig. 2. Ratings for connected and neutral judges, collapsed across story type and target.

higher than connected ones. Separate planned Repeated Measures ANOVAs for the different grades supported this conclusion (kindergartners rate connected judges higher, $F(1, 17) = 13.079$, $p < .01$, $\eta^2 = .435$, second graders show no difference, $F(1, 17) = .553$, $p = .467$.; and fourth, sixth, eighth graders, and adults rate neutral

judges higher, $F(1, 17) = 42.144$, $p < .001$, $\eta^2 = .713$, $F(1, 17) = 149.232$, $p < .001$, $\eta^2 = .898$, $F(1, 17) = 115.726$, $p < .001$, $\eta^2 = .872$, and $F(1, 17) = 98.595$, $p < .001$, $\eta^2 = .853$, respectively).

The nonsignificant difference in judge ratings for second graders raised questions about whether the children were random or patterned in their responses. Out of the 18 second graders, 10 children gave higher ratings of fairness to the neutral judge compared to the connected judge, 6 gave higher ratings to the connected judge, and 2 gave both approximately equal ratings.

As predicted, there was an interaction between contest type and judge rating, in that participants rated the connected judge as less fair for subjective contests than objective ones, $F(1, 102) = 17.144$, $p < .001$, $\eta^2 = .144$. To examine when children recognize that there are some situations in which it is more important to have an impartial judge, we looked at the interaction between judge rating, contest type, and grade (shown in Table 1 and depicted in Fig. 3 with difference ratings between neutral and connected judges for both kinds of contests). This interaction was significant, $F(5, 102) = 3.298$, $p < .01$, $\eta^2 = .139$. In planned Repeated Measures ANOVAs for individual grades, however, we found that the interaction between judge rating and contest type was only significant for eighth graders, $F(1, 17) = 6.267$, $p < .05$, $\eta^2 = .269$, and adults, $F(1, 17) = 21.156$, $p < .001$, $\eta^2 = .554$.

Given that previous research has found that children understand the difference between subjective and objective situations by early elementary school (Mills, 2007), and given that older children and adults are giving higher ratings to judges in objective contests than subjective ones, the nonsignificant findings for children
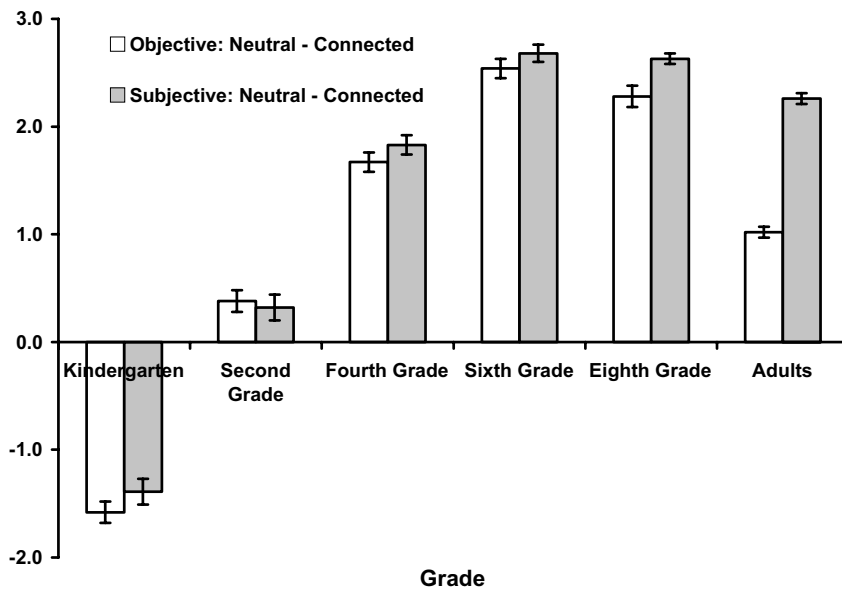

Fig. 3. Difference in ratings between neutral and connected judges for objective and subjective contests.

Table 2
Percentage of participants in each grade showing three patterns of difference in ratings between judges: greater difference for judges in subjective contests, greater difference for judges in objective contests, or no difference

Grade	Percentage of difference in ratings		
	Greater difference for objective contests (%)	No difference (%)	Greater difference for subjective contests (%)
Kindergarten	22	28	50
Second	39	33.5	33.5
Fourth	28	6	66
Sixth	22	28	50
Eighth	6	28	66
Adults	11	11	78

in sixth grade and younger required additional examination. In theory, if children understand the distinction between judges in the two types of contests, there should be a greater difference in ratings for judges in subjective contests compared to judges in objective contests, as found with adults. Therefore, we calculated the difference in ratings for neutral judges compared to connected judges for both types of contests. We then calculated the percentage of participants in each grade who showed a greater difference for subjective contests (the "correct" answer), a greater difference for objective contests, and no difference in ratings across contests, shown in Table 2. The majority of fourth graders, eighth graders, and adults showed a greater distinction between judges in subjective contests than judges in objective contests. Possible explanations for this finding will be reviewed in the discussion.

Finally, there was also a judge by target interaction, in that there were greater differences in ratings for the targets who were connected judges than for the targets who were neutral ones, $F(2, 204) = 9.309$, $p < .001$, $\eta^2 = .084$. The mean ratings for connected judges were 2.716 (friend), 2.820 (parent), and 3.138 (teacher), while the mean ratings for neutral judges were 4.033 (friend), 4.139 (parent), and 4.152 (teacher).

2.3. Discussion

This study addressed two questions. First, do children recognize that personal connections might influence how people judge the outcomes of contests? Our study suggests that there are developmental differences in children's understanding of how personal connections might influence judgments. By fourth grade, and to some extent in second grade, children rated the judge with no personal connection (the "neutral judge") as being more likely to be objective than the judge with a personal connection (the "connected judge"). Kindergartners and some second graders showed the opposite pattern, preferring the connected judge to the neutral one.

Like adults, children also seem to take into account the authority of the potential judge. In this study, we manipulated the judge target to see if the distance from the

child systematically influenced perceptions of objectivity. By second grade, children distinguish between teachers, parents, and friends, concluding that teachers and parents are more likely than friends to be objective in judging contests.

Our second question examined when children come to know that some situations require an impartial judge more than others. Eighth graders and adults clearly distinguished between judges in the objective and subjective contests, rating the connected judge as being less likely to be fair in subjective situations. Given that children as young as kindergarten do have some sense of the distinction between subjective and objective contests (Mills, 2007), it is an open question why younger children are not distinguishing between connected and neutral judges in these two situations, and we will return to this question in the general discussion. Still, when looking at individual patterns of data, the majority of fourth graders, eighth graders, and adults showed a greater distinction between judges in subjective contests than judges in objective contests. Thus, fourth graders may be beginning to understand this distinction.

Sixth graders' failure to make the distinction may be in part because of the method used for testing middle school children and adults. While the paper-and-pencil method has been successful with older children in our research in the past and is more acceptable to their teachers than interviews (see Mills & Keil, 2005), it is more difficult to guarantee that the sixth and eighth graders were attending to the details of each story when making their responses. Future research can involve one-on-one interviews or more detailed fact-checking questions to make sure middle school children are paying attention.

Returning to our first question, young children might prefer the connected judge for a number of reasons. It may be that by asking children to use a rating scale for each individual judge, their understanding of the effects of relationships on judgments is underestimated. That is, the rating scale may have imposed an additional cognitive load that left few reserves for thinking about the influence of connections. Therefore, in Experiment 2, children were presented with pairs of judges and asked to choose which judge they thought would do a better job of picking the legitimate winner.

Children may also think that knowledge and personal connections to the people in the contests are helpful, or even important. Thus, personal connections might provide one with ''inside'' knowledge, thereby allowing a more sophisticated and informed evaluation. In Experiment 2, we examined if young children's preference for a connected judge over a neutral judge is related to having different intuitions than older children about the criteria for being a good judge.

Young children may prefer a judge with a personal connection because the judge would be more aware of a participant's background, regardless of the valence of the connection. In other words, any kind of connection confers an evaluative advantage because it increases factual knowledge about the contestant. Alternatively, given that young children do sometimes use global evaluations to make predictions about future behavior (Alvarez, Ruble, & Bolger, 2001), children might make the inference that someone who has a positive relationship with a contestant is good. Such a judge would be seen as more fair than someone who has no connection (and is thus

neutral). On the other hand, a judge with a negative connection (e.g., enemies with a contestant) would be seen as bad, and such a judge would likely be rated as less fair than a neutral judge.

One way to examine these two alternatives is by presenting children with different types of judges, some connected in a positive way to one contestant (a ''connected friend''), some connected in a negative way (a ''connected enemy''), and some with no connection (a ''neutral party''). If children think judges with any connection are better than judges with none, then they should prefer both connected friends and connected enemies over neutral parties. If, on the other hand, they place value on the valence of the relationship, they should prefer connected friends over neutral parties, as seen in the current study, but they should prefer neutral parties to connected enemies. Experiment 2 examines this question in more detail. Given a general finding in psychological research that negative motives, experiences, and feelings are easier for people to detect and process than positive ones (Baumeister, Bratslavsky, Finkenauer, & Vohs, 2001), we predicted that younger children would prefer the connected friend to the neutral judge, as they did in Experiment 1, but they would prefer a neutral judge to a connected enemy.

3. Experiment 2

Experiment 2 was conducted to address three main questions. First, we wanted to make sure that children would show the same preferences for connected judges compared to neutral ones as seen in Experiment 1 using a forced choice procedure. In order to do this, participants were asked to choose the better of two potential judges for several subjective and objective contests. Second, we aimed to determine if children take into account the valence of the connection of a relationship (whether positive or negative) when determining if the judge will be fair. Therefore, for each contest, we presented children with two items in which a neutral judge was contrasted with either a connected friend or a connected enemy.

Third, we wanted to determine if and when children recognize several other main criteria that make a good judge to see if the performance of younger children in Experiment 1 could be based on them having different intuitions from older children. Adults have long recognized the multiple criteria involved in being a good judge. For example, according to David Hume, a good judge meets the following description: ''strong sense, united to delicate sentiment, improved by practice, perfected by comparison, and cleared of all prejudice, can alone entitle critics to this valuable character,'' (Hume, 1757). As mentioned in the introduction, previous research has found that even young children prefer listening to and learning from speakers who meet some of these criteria, such as being knowledgeable and perceptually capable (e.g., Jaswal & Neely, 2006; Koenig et al., 2004; Sabbagh & Baldwin, 2001). Other research with adults finds that they are concerned about judgments that seem skewed by bias (e.g., Miller & Ratner, 1998).

In piloting with open-ended questions in our laboratory, adult participants discussed many of the same characteristics endorsed by Hume, claiming that a good

judge should be well-informed, practiced at judging similar events in the past, perceptually able to sense the criteria for judging, dispassionate (in other words, not dishonest), without a negative personal connection to a contestant, without a positive personal connection to a contestant, and open-minded. Therefore, in Experiment 2, we contrasted judges who met these criteria for being a good judge with judges who did not, asking children to choose the judge that would do the best job at picking out the person who met the criteria for being the winner of that particular contest (e.g., crossed the finish line first in the running race; was the most beautiful girl in the beauty contest). We predicted that even kindergartners would do well on most of these questions. We based our predictions on prior work showing that young children can take into account knowledge states of others in determining how much to believe them (Jaswal & Neely, 2006; Koenig et al., 2004; Sabbagh & Baldwin, 2001), can recognize that practice improves performance (Lockhart, Chang, & Story, 2002), can understand the importance of perceptual limitations on knowledge (O'Neill, Astington, & Flavell, 1992), and can understand the importance of motivational states on statements (Mills & Keil, 2005). We also predicted that while we would replicate Experiment 1's finding that kindergartners prefer judges with positive connections over neutral ones, we wound find that even kindergartners would prefer neutral judges to judges with negative connections.

3.1. Methods

3.1.1. Participants

Sixteen kindergartners (mean age = 5 years 10 months; range = 5 years 2 months to 6 years 4 months), 16 second graders (mean age = 7 years 8 months; range = 7 years 0 months to 8 years 1 month), and 15 fourth graders (mean age = 9 years 9 months; range = 8 years 6 months to 10 years 2 months). Sixteen adults also participated in the study. The sample was gender balanced and reflected the distribution of ethnic and socioeconomic groups in the community. Children were recruited from the greater New Haven area and were tested in a quiet room; each session took about 15 min.

3.1.2. Design

Participants heard two different kinds of stories: two stories involved objective events (running race, spelling bee), while two involved subjective events (baking contest, art contest). For each contest, participants were asked to choose the better of two potential judges, with one person in the pair meeting one of several criteria for being fair and impartial, while the other person did not meet that particular criterion. The exact wording for this is explained in the section below.

Seven criteria were examined in this experiment. For all of the criteria except open-mindedness, one critical comparison was made for each story between a person who met the criterion and a person who did not, and children were asked to choose the best judge. To ask about the well-informed criteria, children were told about one person who knew a lot about the contest material in question (e.g., ''This person knows a lot about running''), and one person who did not know a lot about that

contest material (e.g., ''This person does not know anything about running.''). For the practiced criteria, a person who had experience judging that kind of contest (e.g., ''This person has judged a running race in the past.'') was contrasted with a person without the experience (e.g., ''This person has never judged a running race before.''). For the perceptually able criteria, a person who was not able to perceive the contest material (e.g., ''This person is blind and cannot see the running race'') was contrasted with a person who could (e.g., ''This person can see.''). For the dispassionate criteria, a person who had cheated at judging a contest in the past was contrasted with a person who had never cheated at judging a contest. For the no personal connections criteria, a neutral judge was contrasted with either a connected friend or a connected enemy.

Finally, for the open-minded criteria, two items were asked that compared a person who held a preexisting bias towards a certain aspect of the contest (e.g., ''This person likes runners who run with their fists closed instead of open'') to a person who was open to multiple methods of competing (e.g., ''This person likes runners who run both with their fists closed and with their fists open.''). Two items were used for each story type to make sure that there was no individual criteria pairing that skewed the results. We also counterbalanced these items across conditions so that in one condition, the person with the preexisting bias preferred one aspect (runners who run with their fists closed instead of open), and in the other condition, the judge held the opposite preference (runners who run with their fists open instead of closed). For the criteria for two sample stories, please see Appendix B.

There were two orders to counterbalance for the different open-minded criteria questions. For those orders, the four contests were read randomly to children. The items were also shuffled so that each child heard the questions about criteria in random order.

3.1.3. Procedure

Similar to Experiment 1, prior to testing, participants were asked if they had ever seen or been in a contest before. The experimenter told the child that, for contests, it is important to have a good judge to determine the winner. They were then told, ''I'm going to tell you about some contests, and I want you to tell me which judge you think would be the most fair. In some situations, it is easier to judge fairly than in other situations. A fair judge should give the prize to the person who really deserves to win: the person who did the best job at the contest they were in. Does that make sense?'' Children were also told that there was no right or wrong answer to the questions. No further training was given. This was to address the possibility that the training from Experiment 1 might have encouraged children to think of only the training examples when determining which judge would be the most fair.

The experimenter then read the stories, periodically asking fact-check questions about the contest. For each of the 4 contests, the same drawings that had been used in Experiment 1 were placed on the table to maintain attention. Following each story, participants heard questions for the different criteria. For each criterion, two judges were presented in random order to the children. A simple line drawing

for each judge was placed on the table in front of the child. Children then had to point to which judge they thought would be the best at picking out who really met the criteria for being the winner of that particular contest (e.g., crossed the finish line first in the running race; made the best drawing in an art contest).

3.2. Results and discussion

To analyze the data, we calculated the number of accurate responses for each criterion (well-informed, experienced, perceptually able to sense the criteria for judging, dispassionate, no negative connection, no positive connection, and open-minded) within each story for participants in each grade. A Repeated Measures ANOVA was conducted for each criterion with story (the two objective stories, running race and spelling bee, and the two subjective contests, baking contest and art contest) as a within-subjects factor and grade as a between-subjects factor. For all but one criterion (open-mindedness), there were no differences in the average number of accurate responses across stories, so for the remainder of the analyses, the data are collapsed across story type, creating a possible score range from 0 to 4. The means are shown in Table 3.

Our main questions that motivated this experiment related to participants' responses on two particular criteria: no positive connection and no negative connection. A one-way ANOVA was computed to examine the responses to the no positive connection question, finding a significant difference between the grades, $F(3, 59) = 31.372$, $p < .001$. Post hoc tests revealed that with the exception of the fourth graders and adults having similar ratings ($M = 3.33$ and $M = 4.00$), the rest of the grades differed from each other. Several follow-up t-tests compared the scores for each grade to the chance level of responding (which would be a score of 2 out of 4). Kindergartners preferred the judge with a positive connection over the one with no connection (the "neutral judge") at above chance levels, $t(15) = 9.586$, $p < .001$, second graders were not significantly different from chance, $t(15) = 0.144$, $p = .889$, and fourth graders and adults preferred the neutral judge, $t(15) = 3.696$, $p < .01$, $t(15) =$ uncalculated (given that the standard deviation was zero). This replicates the findings from Experiment 1 in that kindergartners preferred the judge who was positively biased (i.e., the "connected friend") over the neutral judge, second

Table 3
Means and standard errors for ratings of fairness for each criterion

Grade	No positive connection	No negative connection	Well-informed	Experienced	Perceptually able	Dispassionate
Kindergarten	0.31 (0.18)	3.88 (0.13)	3.69 (0.25)	2.75 (0.45)	3.94 (0.06)	3.94 (0.06)
Second	1.94 (0.43)	4.00 (0.00)	4.00 (0.00)	3.75 (0.11)	3.94 (0.06)	3.94 (0.06)
Fourth	3.33 (0.36)	4.00 (0.00)	4.00 (0.00)	3.73 (0.27)	3.93 (0.07)	3.80 (0.14)
Adults	4.00 (0.00)	4.00 (0.00)	3.94 (0.06)	4.00 (0.00)	4.00 (0.00)	4.00 (0.00)

Note. Mean ratings are based on a 4-point scale; standard errors are in parentheses.

graders showed no preference, and fourth graders and older participants preferred the neutral judge.

The responses to the no negative connection items were very different. A one-way ANOVA found no difference between grades, $F(3, 49) = .978$, $p = .409$. Almost all participants preferred the neutral judge to the judge with a negative connection, and this was at greater than chance levels, $t(62) = 62$, $p < .001$ (see Fig. 4).

Our third main question was to determine if and when children recognize several other main criteria that make a good judge. A one-way ANOVA was conducted for each other criteria (see Fig. 5). For three of the criteria, there was no difference across grade (dispassionate judge who does not cheat, $F(3, 59) = 1.029$, $p = .387$; a judge who can perceive the contest, $F(3, 59) = .342$, $p = .795$; a judge who has knowledge about the contest material, $F(3, 59) = 1.265$, $p = .295$). On average, children in each grade preferred the judges who were well-informed, dispassionate, and had no perceptual disability that could have influenced judgments.

The one-way ANOVA for experience found a significant difference across grades, $F(3, 59) = 4.296$, $p < .01$. In follow-up t-tests, kindergartners did not choose the experienced judge at greater than chance levels, $t(15) = 1.660$, $p = .118$, while second graders and older children did, all ps $< .05$, suggesting developmental differences in the emphasis on judges having experience.

Finally, given that there were differences between stories for the open-minded criterion, that data need to be analyzed slightly differently. The Repeated Measures
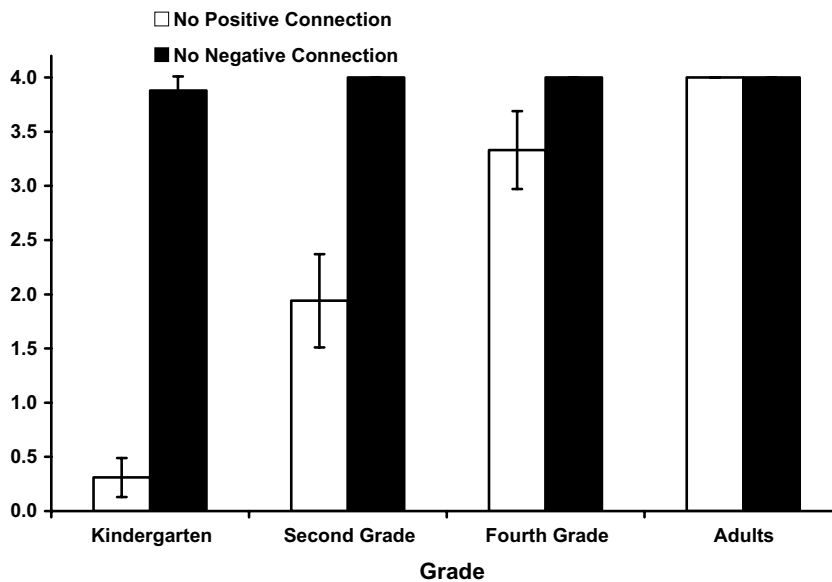

Fig. 4. Mean number of responses (out of 4) per grade showing preference for a judge with no connection compared to a judge with either a positive connection or a negative connection to a contestant.
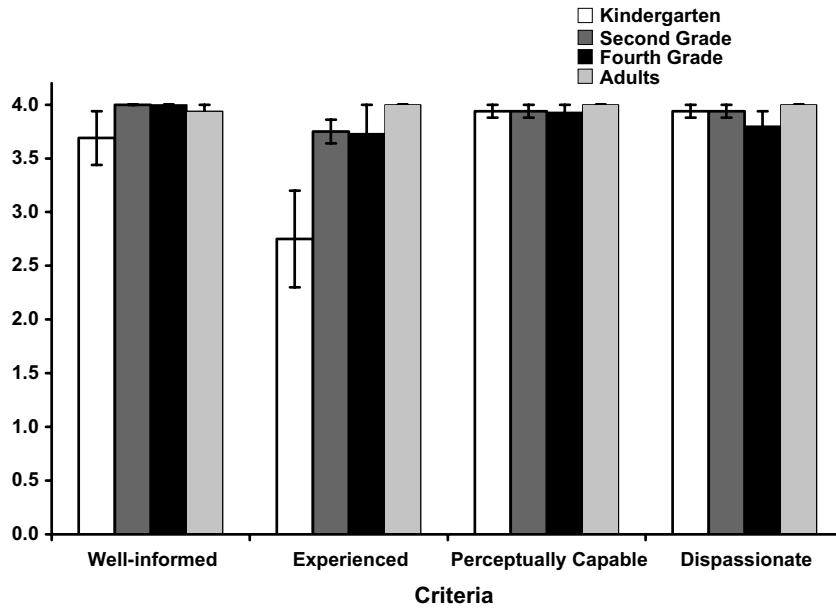
Fig. 5. Mean number preferring a judge meeting four of the criterion across grades.

ANOVA for open-mindedness found a significant difference between stories, $F(7, 413) = 5.183$, $p < .001$, $\eta^2 = .081$. In general, children across grades preferred the judge who is open-minded to a judge who has a specific preference (e.g., prefers small cookies over big ones). A Repeated Measures ANOVA grouping the items from objective stories separately from those in subjective stories found a significant difference between story type, $F(1, 59) = 4.101$, $p < .05$, $\eta^2 = .065$. Participants showed a stronger preference for the open-minded judge in subjective contests ($M = 3.483$) compared to objective contests ($M = 3.323$). For example, if a judge had a preference in an objective contest for runners who finish the race with their head up instead of down, children would often come up with explanations why that could actually be helpful and a good preference to have. If a judge had a preference in a subjective contest, children rejected such explanations. There was also a significant effect of grade, $F(3, 59) = 3.889$, $p < .05$, $\eta^2 = .165$, in that second graders had lower ratings than children in the other grades.

4. General discussion

Overall, these two experiments examined children's understanding of the effects of connections on judgments, the criteria for being a good judge more generally, and the idea that some situations require an impartial judge more than others. Our experiments suggest that children *do* recognize that personal connections might influence

judgments, but that there is a dramatic shift in how they evaluate connected judges. In Experiment 1, by fourth grade, and to some extent in second grade, children rated the judge with no personal connection (the "neutral judge") as being more likely to be objective than the judge with a personal connection (the "connected judge"). Kindergartners and other second graders showed the opposite pattern, preferring the connected judge to the neutral one. This finding was replicated in Experiment 2 when children were asked to choose between a judge who with a positive connection and a judge who was free from positive bias. However, even kindergartners preferred a judge who was free from negative bias to a judge with a negative connection (e.g., enemies).

These findings were not due to kindergartners having a completely different concept of what it means to be a good judge: even kindergartners realized that a good judge should be well-informed, perceptually able to sense the criteria for judging, dispassionate, and open-minded. At the same time, it is clear that kindergartners do think somewhat differently from older children, given that they were much more concerned about connected enemies than connected friends. One possible explanation for this finding is that young children are using global evaluations of a person to determine how much to trust their judgments: friends are "good" and enemies are "bad". It also may be that young children are generally optimistic about people's behavior, assuming that most people are good and will make good choices unless there is clear evidence otherwise. For instance, four to six year olds seem reluctant to make negative trait attributions after hearing about one negative behavior, but they are quick to conclude someone who engages in one positive behavior is "nice", not needing multiple instances to confirm the niceness (Boseovski & Lee, 2006). Another possibility is that children are more attentive to the influence of negative relationships on judgments than positive ones. It has been argued that there is a larger payoff to detecting agents that are malicious than those that are benign, and that humans and other primates employ more cognitive resources in reasoning about malicious agents (Cosmides & Tooby, 2004). Finally, it is also possible that young children assume that people have more useful background knowledge about their friends than about their enemies. It follows that a judge who is friends with a contestant will have more expertise to apply in determining the winner of a contest. Thus, irrespective of the valence of the relationship, children may think the better you know a contestant, the better you would be as a judge. Regardless, the fact that even kindergartners are distrusting of a judge with a negative connection suggests how their first insights into partiality might emerge. Future research conducted by Mills and colleagues will further examine the emergence of an understanding of partiality (Mills, 2006, 2007).

Given the mixed findings for second graders' understanding of positive connections, it would be interesting to determine what distinguishes the children who prefer the neutral judge from those who prefer the one with a positive connection. For example, children who understand that positive connections influence judgments may do better on other measures of advanced social cognition, such as an interpretive theory of mind task (e.g., Carpendale & Chandler, 1996) or a general theory of

mind task (e.g., Wellman & Liu, 2004). There may also be other individual differences, perhaps due to differing desires to evaluate the fairness or accuracy of others. Previous research has found that children of lower socioeconomic status are less likely than children of higher socioeconomic status to rationalize inequities in income across occupations (Emler & Dickinson, 1985), and adults vary in how much they enjoy solving problems and engaging in cognitive endeavors (Cacioppo, Petty, Feinstein, & Jarvis, 1996). Perhaps children who attend more to inequities or like to engage more in cognitive endeavors in general are better at detecting partiality.

The context in which the questions are asked, too, could be a factor. Past research suggests that children's thoughts about the desires of others can vary depending on whether the context invokes moral or personal concerns (Yuill, Perner, Pearson, Peerbhoy, & Van den Ende, 1996). The stories used in the current research involve potential outcomes that influence characters, not the children themselves. It may be that second graders would recognize how personal connections could skew judgments if they were picking judges for a contest they themselves were going to be in, and one of the judges had a personal connection towards someone else (e.g., the mother of the other contestant).

Looking later in development, we examined when children come to know that some situations require an impartial judge more than others. In Experiment 1, eighth graders and adults clearly distinguished between judges in the objective and subjective contests, rating the connected judge as being less likely to be fair in subjective situations. In Experiment 2, kindergartners through fourth graders did not distinguish between connected judges in subjective and objective contexts. However, they showed a stronger preference for the open-minded judges in subjective contexts over objective ones. Their explanations suggested that they thought preexisting biases for criteria in subjective contests (e.g., preferring artists who make big drawings over small drawings) were unfair and inaccurate, while preexisting biases for criteria in objective contests (e.g., preferring runners who run with their fists open instead of closed) were not seen as problematic. One older child pointed out that for the running race example, it did not matter if the judge liked people who ran with their fists open instead of closed, since that would not change who crossed the finish line first. This offers some support that children can distinguish between some aspects of judging in subjective and objective situations.

Given other research findings that even kindergartners do have some sense of the distinction between subjective and objective contests (e.g., Mills, 2007), there are at least two possible reasons why children younger than eighth grade are not distinguishing between connected and neutral judges in these two situations. First, their failure may stem from an assumption that a potentially self-interested judge may make intentionally deceptive judgments. While the likelihood of a judge lying may only vary slightly between objective and subjective situations, the chances of a judge being biased are much more likely for subjective situations, since by definition there is more room in subjective contexts for personal opinion to influence judgments. In this study, eighth graders and adults seem to understand this.

When reasoning about the statements and judgments of others, the fourth and sixth graders in this study may be even more cynical – focused on deceptive intentions – than adults. This interpretation fits with earlier findings showing that younger children do have difficulty understanding biased interpretations made by others, tending to see matters either in terms of malicious lies or truthfulness (Mills & Keil, 2005).

Second, children may have difficulty integrating information about the context with information about the people involved. Children could be focusing primarily on the connection of the judge to the contestants, disregarding information about the context in which the judgment took place. In the current studies, children are explicitly predicting which judge would do a better job, a task that emphasizes judge characteristics rather than context. Thus, one possible direction for future research is to reframe the question to involve explicit predictions of which contest would be harder for a connected judge to make fair judgments. Another possibility is to use an implicit choice task for judges with various characteristics in various contexts: given that previous research in other domains of development finds that children often show implicit understanding before explicit (e.g., Clements, Rustin, & McCallum, 2000), children's performance might improve in this kind of task.

In sum, we find a shift in how children think personal connections might affect judgments. Kindergartners see judges with personal connections as especially fair and just unless they are presented with a judge who likely has negative intentions, while older children are concerned about the potential for deception for both judges with positive connections and those with negative connections. Not until late in middle school do children seem to reflect on aspects of the situation to determine how personal connections might influence reasoning. Thus, children seem to shift from assumptions of beneficence to vigilant guard against malice, before appreciating the relationship between partiality and context. Even having gained such a nuanced understanding as a bystander, when dealing with personally relevant situations such as when considering judges for the highest court in the land, people may be quick to conclude that partiality will lead to unfair or malicious decisions.

Acknowledgements

Appendix A. Experiment 1: sample stories and judges

A.1. Running race

Terence Hall is holding its big annual running race, and Tracy is competing in one of the events. Whoever crosses the finish line first will win a big trophy.

Connected judge: Tracy's mother, who knows many of the kids that are in the race.

Neutral judge: Another mother who does not know any of the people in the race.

A.2. Dance contest

Jessica is in the big county dance contest. The person who does the best dance solo will get a prize.

Connected judge: Jessica's best friend Bethany, who knows many of the girls in the contest.

Neutral judge: Patty, who does not know any of the girls in the contest.

Appendix B. Experiment 2: sample stories

B.1. Running race

Terence Hall is holding its big annual running race. What does the judge have to decide? How can he tell who won?

Let's think about who would do the best job deciding who crossed the finish line first.

1. —— This person is best friends with one of the people in the contest.
   —— This person does not know anyone in the running race.
2. —— This person really does not like one of the people in the contest.
   —— This person does not know anyone in the running race.
3. —— This person likes runners who run with their fists closed instead of open.
   —— This person likes both runners who run with their fists open and those who run with their fists closed.
4. —— This person likes runners who finish with their head down rather than those who finish with their head up.
   —— This person likes both runners who finish with their head down and those who finish with their head up.
5. —— This person cheated when judging a running race last year.
   —— This person has never cheated when judging a running race.
6. —— This person is blind so he cannot see the people run.
   —— This person can see.
7. —— This person knows a lot about running.
   —— This person does not know anything about running.
8. —— This person judged a contest like this last year.
   —— This person has not judged a contest like this before.

B.2. Art

There's an art show at York Middle School, and whoever makes the best artwork will win a prize. People can make drawings or paintings or sketches or whatever they want. What does the judge have to decide? How does he pick which artwork is the best?

Let's think about who would do the best job deciding which artwork is the best.

1. —— This person is best friends with one of the people in the contest.
   —— This person does not know anyone in the art show.
2. —— This person really does not like one of the people in the contest.
   —— This person does not know anyone in the art show.
3. —— This person likes artists who use paints instead of pencils.
   —— This person likes both artists who use paints and those who use pencils.
4. —— This person likes artists who make big drawings more than artists who make small drawings.
   —— This person likes both artists who make big drawings and artists who make small drawings.
5. —— This person cheated when judging an art show last year.
   —— This person has never cheated when judging an art show.
6. —— This person is blind so he cannot see the paintings.
   —— This person can see.
7. —— This person knows a lot about art.
   —— This person does not know anything about art.
8. —— This person judged a contest like this last year.
   —— This person has not judged a contest like this before.

References

Alvarez, J. M., Ruble, D. N., & Bolger, N. (2001). Trait understanding or evaluative reasoning? An analysis of children's behavioral predictions. *Child Development, 72*, 1409–1425.

Baumeister, R. F., Bratslavsky, E., Finkenauer, C., & Vohs, K. D. (2001). Bad is stronger than good. *Review of General Psychology, 5*, 323–370.

Boseovski, J. J., & Lee, K. (2006). Children's use of frequency information for trait categorization and behavioral prediction. *Developmental Psychology, 42*, 500–513.

Bussey, K. (1992). Lying and truthfulness: Children's definitions, standards, and evaluative reactions. *Child Development, 64*, 1005–1021.

Cacioppo, J. T., Petty, R. E., Feinstein, J., & Jarvis, B. (1996). Dispositional differences in cognitive motivation: The life and times of individuals varying in need for cognition. *Psychological Bulletin, 119*, 197–253.

Carpendale, J. I., & Chandler, M. J. (1996). On the distinction between false belief understanding and subscribing to an interpretive theory of mind. *Child Development, 67*, 1686–1706 .

Clements, W. A., Rustin, C. L., & McCallum, S. (2000). Promoting the transition from implicit to explicit understanding: A training study of false belief. *Developmental Science, 3*, 81–92.

Cosmides, L., & Tooby, J. (2004). Social exchange: The evolutionary design of a neurocognitive system. In M. S. Gazzinga (Ed.), *The cognitive neurosciences* (3rd ed., pp. 1295–1308). Cambridge, MA: MIT Press.

Damon, W. (1988). *The moral child*. New York: Free Press.

Emler, N., & Dickinson, J. (1985). Children's representation of economic inequalities: The effects of social class. *British Journal of Developmental Psychology, 3*, 191–198.

Fein, S. (1996). Effects of suspicion on attributional thinking and the correspondence bias. *Journal of Personality and Social Psychology, 70*, 1164–1184.

Flavell, J., Flavell, E., Green, F., & Moses, L. (1990). Young children's understanding of fact beliefs versus value beliefs. *Child Development, 61*, 915–928.

Gnepp, J., & Chilamkurti, C. (1988). Children's use of personality attributions to predict other people's emotional and behavioral reactions. *Child Development, 59*, 743–754.

Heyman, G. D., Fu, G., & Lee, K. (2007). Evaluating claims people make about themselves: The development of skepticism. *Child Development, 78*, 367–375.

Heyman, G. D., & Gelman, S. A. (1999). The use of trait labels in making psychological inferences. *Child Development, 70*, 604–619.

Heyman, G. D., & Legare, C. H. (2005). Children's evaluation of sources of information about traits. *Developmental Psychology, 41*, 636–647.

Hume, D. (1757). Of the standards of taste. In T. H. Green & T. H. Grose (Eds.), *The philosophical works of David Hume* (p. 1874). London: Longman, Green.

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.

Koenig, M. A., Clement, F., & Harris, P. L. (2004). Trust in testimony: Children's use of true and false statements. *Psychological Science, 15*, 694–698.

Kruger, J., & Gilovich, T. (1999). 'Naïve cynicism' in everyday theories of responsibility assessment: On biased assumptions of bias. *Journal of Personality and Social Psychology, 74*, 743–753.

Kuhn, D., Cheney, R., & Weinstock, M. (2000). The development of epistemological understanding. *Cognitive Development, 15*, 309–328.

Kuhn, D., & Weinstock, M. (2002). What is epistemological thinking? In B. Hofer & P. Pintrich (Eds.), *Personal epistemology: The psychology of beliefs about knowledge and knowing*. Mahwah, NJ: Erlbaum.

Lampinen, J. M., & Smith, V. L. (1995). The incredible (and sometimes incredulous) child witness: Child eyewitnesses' sensitivity to source credibility cues. *Journal of Applied Psychology, 80*, 621–627.

Lee, K., Cameron, C. A., & Doucette, J. (2002). Phantoms and fabrications: Young children's detection of implausible lies. *Child Development, 73*, 1688–1702.

Lerner, M. J. (1974). The justice motive: 'Equity' and 'parity' among children. *Journal of Personality and Social Psychology, 29*, 539–550.

Lockhart, K. L., Chang, B., & Story, T. (2002). Young children's beliefs about the stability of traits: Protective optimism? *Child Development, 73*, 1408–1430.

Lutz, D. R., & Keil, F. C. (2002). Early understanding of the division of cognitive labor. *Child Development, 73*, 1073–1084.

McGraw, K. M., Lodge, M., & Jones, J. (2002). The pandering politicians of suspicious minds. *The Journal of Politics, 62*, 362–383.

Miller, D. T., & Ratner, R. K. (1998). The disparity between the actual and assumed power of self-interest. *Journal of Personality and Social Psychology, 74*, 53–62.

Mills, C. M. (2006). Children's detection of (im)partiality. Poster presented at the 2006 Association for Psychological Science 18th Annual Convention, New York, NY.

Mills, C. M. (2007). Understanding the effects of context and relationships on judgments. In C. M. Mills (Chair), *Taking a critical stance: How children evaluate the thinking of others*. Symposium conducted at the Biennial Meeting of the Society for Research in Child Development, Boston, MA.

Mills, C. M., & Keil, F. C. (2005). The development of cynicism. *Psychological Science, 16*, 385–390.

Nelson, S. A., & Dweck, C. S. (1977). Motivation and competence as determinants of young children's reward allocation. *Developmental Psychology, 13*, 192–197.

Nichols, S., & Folds-Bennett, T. (2003). Are children moral objectivists? Children's judgments about moral and response-dependent properties. *Cognition, 90*, B23–B32.

O'Neill, D. K., Astington, J. W., & Flavell, J. H. (1992). Young children's understanding of the role that sensory experiences play in knowledge acquisition. *Child Development, 63*, 474–490.

Pasquini, E. S., Corriveau, K. H., Koenig, M., & Harris, P. (2007). Preschoolers monitor the relative accuracy of informants. *Developmental Psychology, 43*, 1216–1226.

Petty, R. E., & Cacioppo, J. T. (1986). *Communication and persuasion: Central and peripheral routes to attitude change*. New York: Springer.

Pillow, B. H. (1991). Children's understanding of biased social cognition. *Developmental Psychology, 27*, 539–551.

Rawls, J. (1971). *A theory of justice*. Cambridge, MA: Harvard University Press.

Robinson, E. J., Mitchell, P., & Nye, R. (1995). Young children's treating of utterances as unreliable sources of knowledge. *Journal of Child Language, 22*, 663–685.

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.

Sigelman, C. K., & Waitzman, K. A. (1991). The development of distributive justice orientations: Contextual influences on children's resource allocations. *Child Development, 62*, 1367–1378.

Sodian, B. (1988). Children's attributions of knowledge to the listener in a referential communication task. *Child Development, 59*, 378–385.

Thierry, K. L., & Spence, M. J. (2002). Source-monitoring training facilitates preschoolers' eyewitness memory performance. *Developmental Psychology, 38*, 428–437.

Thierry, K. L., Spence, M. J., & Memon, A. (2001). Before misinformation is encountered: Source monitoring decreases child witness suggestibility. *Journal of Cognition and Development, 2*, 1–26.

Thorkildsen, T. A., & White-McNulty, L. (2002). Developing conceptions of fair contest procedures and the understanding of skill and luck. *Journal of Educational Psychology, 94*, 316–326.

Welch-Ross, M. (1999). Interviewer knowledge and preschoolers' reasoning about knowledge states moderate suggestibility. *Cognitive Development, 14*, 423–442.

Wellman, H., & Liu, D. (2004). Scaling of theory-of-mind tasks. *Child Development, 75*, 523–541.

Wimmer, H., & Perner, J. (1983). Beliefs about beliefs: Representation and constraining function of wrong beliefs in young children's understanding of deception. *Cognition, 13*, 103–128.

Yuill, N., & Pearson, A. (1998). The development of bases for trait attribution: Children's understanding of traits as causal mechanisms based on desire. *Developmental Psychology, 34*, 574–586.

Yuill, N., Perner, J., Pearson, A., Peerbhoy, D., & Van den Ende, J. (1996). Children's changing understanding of wicked desires: From objective to subjective to moral. *British Journal of Developmental Psychology, 14*, 457–475.