Available online at www.sciencedirect.com
SCIENCE DIRECT®
Cognitive Development 19 (2004) 111–125
COGNITIVE DEVELOPMENT

The relative power of an emotion's facial expression, label, and behavioral consequence to evoke preschoolers' knowledge of its cause

Sherri C. Widen, James A. Russell*

Department of Psychology, McGuinn Hall, Boston College, 140 Commonwealth Avenue, Chestnut Hill, MA 02467, USA

Received 1 February 2003; received in revised form 31 August 2003; accepted 1 November 2003

Abstract

Lay people and scientists alike assume that, especially for young children, facial expressions are a strong cue to another's emotion. We report a study in which children ($N = 120$; 3–4 years) described events that would cause basic emotions (surprise, fear, anger, disgust, sadness) presented as its facial expression, as its label, or as its behavioral consequence. For no emotion was the facial expression the strongest cue. Performance for fear and disgust was more accurate given its label or its behavioral consequence than given its facial expression; performance for anger was more accurate given the consequence. For 3s, behavioral consequences were the strongest cues to emotion; for 4s, labels were.

Keywords: Facial expression; Preschooler; Performance

By figuring out that Dad is angry, Chris can learn more about anger: What causes it, what its consequences are, how it is expressed, how it is regulated, how it is labeled, and so on. Acquiring this information is important for various aspects of a child's social functioning, such as honing social skills, forming friendships, developing positive peer relations, and adjusting to school (Denham, 1998; Shields et al., 2001; Smith, 2001). Theorists have suggested that, even earlier in life, an understanding of emotion paves the way for infant-caregiver attachment (Bowlby, 1969, 1988; De Rosnay & Harris, 2002; Magai & McFadden, 1995), which has in turn been implicated in preschoolers' cognitive and linguistic development (Robinson & Acevedo, 2001). Acquisition of emotion knowledge is clearly important in

* Corresponding author. Tel.: +1-617-552-4546; fax: +1-617-552-0523.
E-mail address: james.russell@bc.edu (J.A. Russell).

0885-2014/$ – see front matter
doi:10.1016/j.cogdev.2003.11.004

the development of emotional intelligence (Barrett & Salovey, 2002). But, how does the process get started? How did Chris know that Dad was angry in the first place?

Two broad approaches to emotion suggest different answers. Those who view emotions as biologically given response modules have emphasized *facial expressions* as inherited signals that are easily or even innately recognized (Darwin, 1965; Izard, 1971, 1994). This line of thought, of major importance in the psychology of emotion, suggests that Dad's angry facial expression signaled his anger to Chris, who recognized that signal, and thereby began the process of learning about anger. Theoretical considerations of the evolutionary advantages of an emotion signaling system between infant and caregiver (Bowlby, 1969, 1988; Denham, 1998; Izard, 1971; Harris, 1989) have been taken to suggest that very young children recognize specific "basic" emotions from facial expressions. Harris (1989) proposed that an early understanding of facial expressions leads to an understanding of other aspects of emotions, which leads in turn to a theory of mind. Denham (1998) pointed to this early understanding of emotion via facial expressions as the "perceptual bedrock" (p. 61) for all later understanding of emotion.

Various sources of evidence can be cited as support for the theory that even very young children understand facial expressions. Infants distinguish different facial expressions and respond appropriately to them (Barrera & Maurer, 1981; Caron, Caron, & MacLean, 1988; Klinnert, Emde, Butterfield, & Campos, 1988; Maurer & Barrera, 1981; Moses, Baldwin, Rosicky, & Tidball, 2001; Serrano, Iglesias, & Loeches, 1992, 1995; Walker-Andrews & Lennon, 1991). There is also evidence that preschoolers can label facial expressions of the basic emotions (happiness, sadness, anger, fear) with above-chance accuracy (Denham & Couchoud, 1990; Harrigan, 1984; Markham & Adams, 1992; Russell & Widen, 2002a; Widen & Russell, 2003a). When preschoolers are presented with a conflicting facial expression and situation (e.g., someone receiving a gift but displaying a sad facial expression), and are asked to label the protagonist's emotion, they focus more on the facial expression than on the situation (Gnepp, 1983; Reichenbach & Masters, 1983; Wiggers & van Lieshout, 1985). Young preschoolers also label facial expressions more accurately than brief stories describing the same emotion's causes and consequences (Widen & Russell, 2003b).

From a very different perspective, those who think of emotions as social constructions have focused on *emotion labels* as central to a person's understanding of emotion (e.g., Harré, 1986). This line of thought suggests that someone's labeling Dad as *angry* might be what started Chris's process of forming a concept of anger and thus learning about it. Linguistic evidence on cultural differences in emotion concepts can be cited as evidence for this perspective (Harré, 1986; Lutz, 1988; Wierzbicka, 1992). For preschoolers, there is some evidence for what has been called a Label Superiority Effect. Children (preschool to second grade) heard stories about emotional events and selected the protagonist's emotion from an array of three facial expressions or from an array of three labels (Camras & Allison, 1985). Surprisingly, children were more accurate with labels than with facial expressions, especially for fear and disgust. When children (4–5 years) were presented with a facial expression or a label and then asked to describe the emotion's cause (Russell, 1990; Russell & Widen, 2002b), labels were again stronger cues than facial expressions overall, especially for fear and disgust. This advantage of labels over facial expressions also occurred for 2- to 7-year-olds for happiness, sadness, and anger in a sorting task (Russell & Widen, 2002a): Children sorted facial expressions more accurately given the emotion label than

the corresponding facial expression. Thus, the Label Superiority Effect has proven robust to several changes in method.

There is another obvious possibility, namely that children attend to overt *behavior* as a cue to the other's emotion. Children find behavior highly salient and use their observations of another's behavior to solve problems (Call & Tomasello, 1995; Want & Harris, 2001) and to learn social rules (Bandura, 1992) and gender roles (Powlishta, Sen, Serbin, Poulin-Dubois, & Eichstedt, 2001). They seek to understand, predict, label, and influence the behavior of others. Some evidence shows that young children associate emotions with behavioral consequences (Heyman & Gelman, 1998; Russell, 1990; Surbey, 1979). Surbey showed that 5-, 7-, and 9-year-olds can generate the behavioral consequences of basic emotions given the emotion's label. Russell showed that 4- and 5-year-olds can generate accurate causes and consequences, given either a label or a facial expression for a basic emotion. This line of thought suggests that Dad's hostile behavior might be what got Chris's process of understanding anger started. From this perspective, the behavioral consequences of emotions are strong cues to emotions. To our knowledge, no prior research compared the power of behavioral consequences to other cues to emotion.

In the study reported here, we asked children to imagine events that would cause happiness, sadness, anger, fear, surprise, and disgust. The emotion was presented as its facial expression, its label, or its behavioral consequence. The power of a given cue in evoking a child's causal knowledge of an emotion is one indication of which cue plays a larger role in acquisition of that knowledge. The primary purpose of this study was to examine the power of behavior (or wishes to behave)—relative to facial expressions and labels—in eliciting children's knowledge of the emotions' causes. The behavioral consequences were brief descriptions of stereotypical behavioral and observable physiological responses to emotional events (e.g., for fear, screaming and wanting to run away; for sadness, crying) (see Table 1). The 'biological' bedrock hypothesis suggests facial expressions will be stronger cues than behavioral consequences. Social construction theory implies that labels will be stronger cues than behavioral consequence. However, we suggest, behavioral consequences are potentially as powerful cues to emotion as the others.

Table 1
Behavioral consequence used for each emotion

Sadness	. . .	while D was in the living room, something happened that made him feel a certain way. D walked slowly over to a chair and sat down. Tears came to his eyes. He didn't want to talk to anyone.
Anger	. . .	while D was at school, something happened that made him feel a certain way. It made D yell and hit another kid. He clenched his fist and stomped his feet. He yelled really loud.
Fear	. . .	while D was at the park, something happened to D that made him feel a certain way. It made D scream. He ran away as fast as he could. D kept looking back to see if he was being followed. He just wanted to get home where he was safe.
Surprise	. . .	while D was outside, something happened that made him feel a certain way. It made D stop and stand completely still. His heart was beating very fast. He didn't know what had happened. He looked around and tried to figure out what it was.
Disgust	. . .	while D was in the kitchen, something happened that made him feel a certain way. It made D want to wash. He wanted to get it off of himself as fast as he could. He didn't want to touch that stuff.

Note. Each story began with, "The next day . . . ." 'D' stands for 'David' or 'Danny,' which ever name the child chose for the protagonist.

A secondary purpose of this study was to examine the Label Superiority Effect in 3s (and, for comparison, 4s). The Label Superiority Effect has been observed mainly in older preschoolers (4s and 5s in Widen & Russell, 2002; 'preschoolers' in Camras & Allison, 1985), but results were mixed for 3s: Although it has also been reported once for 3s (Russell & Widen, 2002b), the opposite was found in another study—facial expressions were stronger cues than labels—for younger preschoolers (2s and 3s in Widen & Russell, 2003b). Clearly for even younger children, who have not yet acquired emotion labels, facial expressions are the more powerful cue. Therefore, 3 years of age may be a transitional period, providing the toughest test of the Label Superiority Effect.

1. Method

1.1. Participants

Participants were 120 children, all proficient in English and enrolled in daycares in the Greater Boston, MA, area. There were 30 boys and 30 girls in each of two age groups: 3s (35–47 months; mean = 42 months) and 4s (48–70 months; mean = 58 months).

1.2. Materials

1.2.1. Photographs of facial expressions

There were six $5'' \times 7''$ black-and-white glossy photographs of posed prototypical facial expressions posed by a 12-year-old boy. The photographs were provided by Dr. Linda Camras. Camras, Grow, and Ribordy (1983) described the development of the photographs, their coding according to Ekman and Friesen's (1978) Facial Action Coding System, and their use in a study on recognition of facial expressions.

1.2.2. Labels

The labels for the six emotions were *happy*, *sad*, *angry*, *scared*, *surprised*, and *disgusted*.

1.2.3. Behavioral consequence stories

Six stories describing stereotypical behavioral consequences of each emotion were created (Table 1) based on prior work in our lab in which slightly older children generated consequences of specific emotions (Russell, 1990).

1.3. Design and procedure

The experimenter (one of two women) began by spending time playing and chatting with each child until the child seemed comfortable with the experimenter. On a later visit, the experimenter invited the child to play a story-telling game. The experimenter introduced the game: "In this game, we are going to take turns telling a story about things that happen to a boy. The boy is 8 years old. You get to have your turn first. You get to pick his name. Do you want to name him Davie or Danny? [Pauses for child's response] Now, it is my turn. I think that I'll say: D lives with his mom and his dad. Okay, it is your turn. What kind of pet should

D have? [pauses] And what color is it? [pauses] Is it a great big one or a little tiny one? [pauses] So, D is 8 years old, and he lives with his mom and his dad. And D has a pet ___ ."

The happy trial was always first. The experimenter began: "Now, let's make up a story about something that happens to D. Remember, we are just making up a story. So, when it is your turn, you can make up anything you want. Okay? Ready? I'll take my turn first. Once upon a time, there was a boy named D. One day, while D was in his bedroom [showing illustration] something happened that made him feel a certain way. It made D feel happy. It made D do a dance and clap his hands. It made D look like this [showing happiness facial expression]. Now it is your turn. What happened? What made D feel like that?" On this and all subsequent emotion trials, the experimenter used the same, neutral tone of voice.

Up to this point, all the children had been treated identically. Children were now randomly assigned to one of three modes of presentation—facial expression, label, behavioral consequence—which was thus a between-subjects condition. In each mode of presentation, the experimenter continued seamlessly on to the next emotion trial, but now the child was presented with only the facial expression, or the label, or the consequence (instead of all three as in the happy trial). The experimenter said, "Do you want to make up another story? [pauses for child's consent] The next day, when D was [identify setting: e.g., at the park], something happened that made D feel a certain way . . . ." The experimenter introduced the next emotion, selected randomly, in the appropriate mode of presentation. On each trial (in all modes of presentation), the child was shown a different simple illustration of the setting of the story for each emotion: for sadness, a living room; for anger, a school classroom; for fear, a sandbox; for surprise, a backyard; and for disgust, a kitchen. The five emotion trials (sadness, anger, fear, surprise, disgust) were randomly ordered. When each trial ended, the experimenter praised the child and the next emotion trial was introduced with, "Do you want to do another one? [pause for child's consent] The next day . . . ."

1.3.1. Facial expression mode of presentation

In the facial expression mode of presentation, each emotion was presented as its prototypical facial expression. On each trial, the experimenter said, "It made him feel like this [pointing to a photograph of the facial expression]. D was feeling so much like this [pointing] that his mom could tell, his dad could tell, and all his friends could tell that D was feeling like this [pointing]." The facial expression was visible throughout the trial.

1.3.2. Label mode of presentation

In the label mode of presentation, each emotion was presented orally as its label. On each trial, the experimenter said, "It made him feel very, very [label]. D was feeling so [label] that his mom could tell, his dad could tell, and all his friends could tell that D was feeling [label]."

1.3.3. Consequence mode of presentation

In the consequence mode of presentation, each emotion was presented as a brief description of the behavioral consequence of that emotion (see Table 1).

1.3.4. Dependent measure

On each trial (in all modes of presentation), after the emotion was presented, the experimenter asked, "What happened? What made D feel like that?" If the response was a

"non-story," the experimenter prompted the child. (A "non-story" was a response devoid of information about why D would have any emotion. For example, the child did not respond, said "I don't know," or gave an irrelevant response such as "The flower," "I can't think of anything" or "What is that facial expression?") The first prompt was repeating the question (What made D feel like that?). The second prompt was to ask if someone was there with D, and what that person could do to make D feel that way. The third was "What would make *you* feel . . . " and completing the question with "this way" while pointing to the facial expression in the face mode, with the emotion label in the label mode, and with the story (using the child as the protagonist) in the story mode. If the child still did not respond to this last prompt, the child's response was scored as a "non-story": The experimenter then completed the story with a pre-designated ending (e.g., for happy: "His mom made him chocolate chip cookies; and that made him feel like that." For sad: "He fell down and hurt his knee, and that made him feel like that").

1.4. Scoring of responses

The happy trial served as a screening device: The child had to produce a plausible cause for happiness to be included in the sample. A total of eight children (four 3-year-old boys, two 4-year-old boys; two 3-year-old girls) were excluded because their stories were not plausible for happiness (e.g., "His mom ate worms"; "The drawers, because they are blue"); an additional 10 children (five 3-year-old boys, one 4-year-old boy; four 3-year-old girls) did not complete all six trials. These children were replaced by same-sex age-mates. Thus, all children included in the sample showed that they understood, and were willing to complete, the story-telling task. Responses to the happy trial were not included in subsequent analyses.

Collectively, the children had 600 opportunities to tell a causal story. Of these, 66 were non-stories (21 in the face mode, 23 in the label mode, 22 in the story mode). The remaining 534 were stories, which were read to three raters (blind to mode of presentation, the child's age and sex, and the target emotion), who made two judgments: (a) their best guess as to which of the six emotions the child was responding and (b) (no longer blind to the target emotion) a yes/no judgment on the plausibility of the child's story for the actual emotion to which the child was responding. For each rating, each rater made her judgment independently. For those stories on which the three raters did not agree as to specific emotion or plausibility, consensus was reached by discussion.

1.4.1. Best guess

For the best-guess procedure, the rater chose from a list (happiness, surprise, fear, disgust, anger, sadness) the emotion that best suited the child's story. Prior to discussion, all three raters agreed on a specific emotion for 71.8% of the stories (chance = 4.6%; Comparison of proportions, two-sided: $P < 0.001$). Examples of stories scored as correct by the best-guess criterion are: for happiness, "His friend gave him presents for Christmas"; for sadness, "He wanted to go outside but it was raining"; for anger, "Someone took something away from him"; for fear, "There was a monster behind him with a lot of eyes"; for surprise, "His mom came out and she was bigger than his dad, even bigger than the house"; for disgust, "He ate rotten milk."

1.4.2. Plausibility ratings

For the plausibility rating, the three raters were first given the correct emotion label, and each was then asked independently to decide whether the story was a plausible cause of that emotion. Prior to discussion, all three raters agreed on whether it was plausible for 89.3% of the stories (chance = 12.5%; Comparison of proportions, two sided: $P < 0.001$). Examples of stories for which the best guess was incorrect, but the story was rated as plausible are: "If somebody took something of mine," which was generated and rated as plausible for *sad*, but for which the best-guess rating was *angry*; "His daddy came home," which was generated and rated as plausible for *surprised*, but for which the best-guess rating was *happy*.

2. Results and discussion

Of the 600 responses, the 66 "non-stories" were automatically scored as incorrect. Of the remaining 534 stories, the raters' best guess was correct for 242 (45.1%) (chance = 16.7%); and they judged an additional 139 to be plausible. Thus, of the total 600 responses, 40.1% were scored as correct by the best-guess criterion, 63.5% by the plausibility criterion. The number of non-stories and large number of incorrect responses show that the task of generating plausible causes is not easy for this age group and would be difficult to use with younger children. (Similar results were reported by Russell & Widen, 2002b, for a similar task.) Russell (1990) argued that the best-guess and plausibility ratings estimate upper and lower bounds, respectively, of children's knowledge. Convergence of results across the two measures is the best indication of a valid finding.

In two parallel repeated-measures ANOVAs ($\alpha = 0.05$), mode of presentation (3 levels: facial expression, label, consequence), age (two levels: 3s, 4s), and sex (two levels) were between-subjects factors; and emotion (sadness, anger, fear, surprise, disgust) was a within-subject factor. The dependent variable was whether the child's causal story was correct or not, scored 1 or 0. In the first analysis, correctness was determined by the best-guess criterion; in the second, by the plausibility criterion.

Some results were standard: First, in both analyses, there were no main or interactive effects for sex. Second, in both analyses, there was a main effect for emotion: for best guess, $F(4, 432) = 5.38$, $P < 0.001$; for plausibility, $F(4, 432) = 8.47$, $P < 0.001$. As shown in Table 2, sadness, anger and fear were generally easier for children than surprise or disgust. Finally, there was a main effect for age, although it was significant only in the plausibility analysis, $F(1, 108) = 4.38$, $P = 0.04$—proportion correct was 0.68 for the 4s versus 0.59 for the 3s. In the best-guess analysis, the 4s' (0.43) performance was higher than the 3s' (0.37), but this difference was not significant ($P = 0.15$).

Of greater interest was the effect of mode of presentation. There was a significant main effect for mode (Table 2): for best guess, $F(2, 108) = 7.77$, $P < 0.001$; for plausibility, $F(2, 108) = 3.85$, $P = 0.02$. In both analyses, performance in the label mode was significantly (Least Significant Difference [LSD] comparisons; $P < 0.01$) higher than in the facial expression mode. In neither analysis did performance in the label and consequence modes differ significantly. By the best-guess criterion (but not the plausibility criterion), performance in the consequence mode was significantly ($P = 0.004$) higher than in the facial expression mode.

Table 2
Effect of mode of presentation on correctness of causal stories, separately by the best-guess criterion and plausibility criterion

Emotion	Mode of presentation			Mean
	Facial expression	Label	Behavioral consequence	
Proportion correct by best-guess criterion				
Sadness	0.43_{ac}	0.58_{ab}	0.50_{ae}	0.50_{r}
Anger	0.35_{acd}	0.38_{ace}	0.58_{e}	0.43_{rt}
Fear	0.25_{acd}	0.65_{b}	0.48_{be}	0.46_{rt}
Surprise	0.28_{cd}	0.30_{c}	0.18_{c}	0.25_{s}
Disgust	0.15_{d}	0.50_{ac}	0.48_{a}	0.38_{t}
Mean	0.29	0.48	0.44	
Proportion correct by plausibility criterion				
Sadness	0.70_{ac}	0.80_{a}	0.73_{a}	0.74_{r}
Anger	0.83_{a}	0.75_{a}	0.75_{a}	0.78_{r}
Fear	0.53_{c}	0.68_{ac}	0.58_{ac}	0.59_{s}
Surprise	0.58_{c}	0.68_{ac}	0.48_{c}	0.58_{s}
Disgust	0.23_{b}	0.63_{a}	0.63_{ac}	0.49_{s}
Mean	0.57	0.71	0.63	

Note. $\alpha = 0.05$ for all Fisher's Least Significant Differences (LSD) comparisons. Post hoc analyses were calculated separately for each criterion; comparisons between criteria were not done. For the best-guess criterion, means in the same row that do not share a subscript differ significantly, $P \leq 0.05$; means in the same column that do not share a subscript differ significantly, $P \leq 0.05$. Means for emotion that do not share a subscript differ at $P < 0.04$. For the plausibility criterion, means in the same row that do not share a subscript differ significantly, $P < 0.001$; means in the same column that do not share a subscript differ significantly, $P \leq 0.01$. Means for emotion that do not share a subscript differ at $P < 0.01$.

The main effects of age and mode of presentation must be qualified by a significant age × mode-of-presentation interaction (Table 3): for best guess, $F(2, 108) = 4.42$, $P = 0.01$; for plausibility, $F(2, 108) = 4.71$, $P = 0.01$. In both analyses, LSD comparisons showed that with age performance improved significantly ($P < 0.01$) in the label mode only. Put differently, at both ages, facial expression was the least powerful cue and there was no significant increase in its power with age. The most powerful cue to emotion varied with age: For the 3s, the consequence mode was highest by both criteria, followed by the label, then the facial expression; the difference between consequence and facial expression modes was significant ($P = 0.004$) for the best-guess criterion only. For the 4s, the label mode was significantly higher than either of the other modes ($P \leq 0.01$); performance in the consequence and facial expression modes did not differ significantly.

The mode-of-presentation × emotion interaction was also significant in both analyses (Table 2): for best guess, $F(8, 432) = 2.55$, $P = 0.01$; for plausibility, $F(8, 432) = 2.52$, $P = 0.01$. For sadness, fear, and disgust, results fit the pattern described for the main effect of mode: label was the best cue, followed by consequence, followed by facial expression. (The advantage of the label and consequence modes over the facial expression mode was significant for disgust (LSD; $P < 0.01$) by both criteria, and for fear ($P \leq 0.05$) by the best-guess criterion.) For anger and surprise, in contrast, a different pattern occurred. For anger, consequence was the best cue—significantly ($P \leq 0.05$) higher than the facial

Table 3
Effect of age and mode of presentation

Age	Mode of presentation			Mean
	Facial expression	Label	Behavioral consequence	
Proportion correct by best-guess criterion				
3 years	0.27_{a}	0.37_{ab}	0.48_{b}	0.37
4 years	0.31_{a}	0.59_{c}	0.40_{ab}	0.43
Mean	0.29	0.48	0.44	
Proportion correct by plausibility criterion				
3 years	0.55_{a}	0.58_{a}	0.65_{a}	0.59
4 years	0.59_{a}	0.83_{b}	0.61_{a}	0.68
Mean	0.57	0.71	0.63	

Note. $\alpha = 0.05$ for all Fisher's LSD comparisons. Post hoc analyses were calculated separately for each criterion; comparisons between criteria were not done. For the best-guess criterion, means in the same row that do not share a subscript differ significantly, $P < 0.01$; means in the same column that do not share a subscript differ significantly, $P \leq 0.003$. For the plausibility criterion, means in the same row that do not share a subscript differ significantly, $P < 0.002$; means in the same column that do not share a subscript differ significantly, $P \leq 0.001$.

expression mode and marginally higher than the label mode ($P = 0.08$). There were no significant differences among modes for surprise.

Because mode of presentation interacted separately with age and with emotion, we calculated ANOVAs ($\alpha = 0.05$) separately for each age. These analyses, which are illustrated in Fig. 1, replicated the effect of mode within each age group: for 3s, $F(2, 57) = 4.14$, $P = 0.02$; for 4s, $F(2, 57) = 8.10$, $P < 0.001$. These analyses also clarified the mode × emotion interaction. For anger, the effect of mode was significant only for 3s—$F(2, 57) = 5.05$, $P = 0.01$—for whom performance in the consequence mode (0.65) was significantly ($P = 0.01$) higher than in the label (0.25) and facial expression (0.25) modes. For fear, effect of mode was significant only for the 4s, $F(2, 57) = 5.70$, $P = 0.01$; performance in the label mode (0.75) was significantly higher ($P = 0.001$) than in the facial expression mode (0.25); performance in the consequence mode (0.50) did not differ significantly from the other two modes. For disgust, effect of mode was also significant only for the 4s, $F(2, 57) = 6.36$, $P = 0.003$; performance in the label (0.50) and consequence (0.45) modes was significantly higher ($P < 0.01$) than in the facial expression mode (0.05). To summarize, significant differences observed in the mode × emotion interaction did not replicate across age groups. Rather, for 3s, the consequence was the superior cue for anger; for 4s' labels were the superior cue for fear and disgust.

Previous studies (Bullock & Russell, 1985, 1986; Russell, 1990; Russell & Widen, 2002b) had found that children's implicit definition of the label *surprised* differed from their implicit definition of the facial expression of surprise. The label *surprised* was more like the adult definition of *excited* or *happy*—a child's sense of the word *surprised* has a positive valence. In contrast, their implicit definition of the surprise facial expression is more neutral in valence. The present results replicated and extended this effect (Table 4). As expected, the causes produced for the surprise *label* were positive (0.75) significantly (comparison of proportions, two-sided; $P = 0.001$) more often than negative (0.25). And causes for

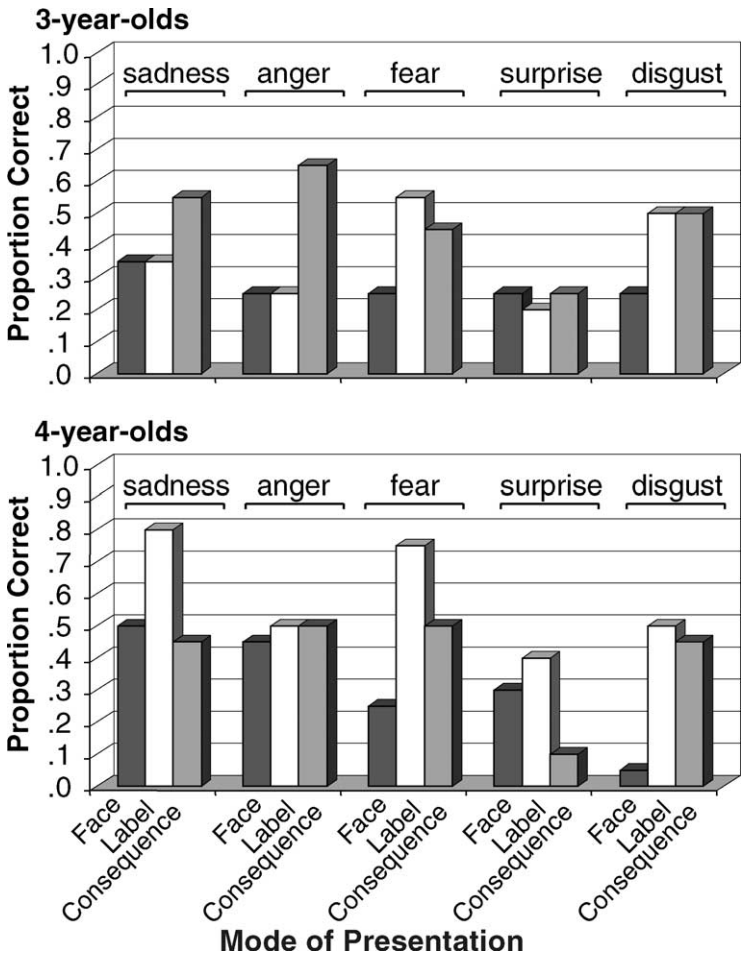

Fig. 1. Effect of mode of presentation for each emotion for the best-guess criterion for 3- and 4-year-olds.

Table 4
Surprise stories rated as positive or negative in each mode of presentation

Rating	Mode of presentation		
	Facial expression	Label	Behavioral consequence
Positive	11	18	3
Negative	8	6	18

Note. $N = 64$. All stories in response to the emotion surprise that were incorrect by the best-guess criterion were analyzed. Those stories for which the raters' best guess was *happy* were classified as positive; those stories for which the best guess was anger, fear, sad, or disgust were classified as negative. A Chi-square comparison of the label and consequence modes was significant, $\chi^2_{df=1} = 16.59$, $P < 0.001$, as was the comparison of the facial expression and consequence modes, $\chi^2_{df=1} = 8.34$, $P = 0.004$.

the surprise *facial expression* were positive (0.58) only slightly more often than negative (0.42) (comparison of proportions was n.s., $P = 0.33$). Interestingly, causes for the surprise *consequence*, in contrast, were negative (0.86) significantly ($P = 0.02$) more often than positive (0.14). No other emotion showed this pattern of valence differing with mode of presentation. These results hint that different cues do not tap the identical concept of surprise. If replicable, they may indicate that children are remembering specific instances, and have yet to join these instances into a single concept. If so, surprise would conform to instance-based, or exemplar, models of concept formation.

3. Conclusion

In this study, 3- and 4-year-olds were asked to imagine what might have caused another child's emotion. The task per se was within their abilities: 100% generated a plausible causal story for happiness. For the other five emotions, although the children generated stories 89% of the time, the stories were correct only about half the time (43% by best guess, 63% by plausibility). What was interesting was that their performance describing a cause varied with the cue to the other's emotion.

Showing children the hypothesized universal, biologically based facial expression for a basic emotion was overall the least good way to specify that emotion. Our results thus challenge the assumption that prototypical facial expressions are especially potent cues to emotion for 3- and 4-year-olds. At this age, labels or behavioral consequences were both stronger cues overall. Although the facial expression mode was relatively weak in this study, we are not willing to abandon the notion that facial expressions have a role to play in the development of children's understanding of emotion, although perhaps a more limited role than sometimes envisioned. Especially for children younger than 3 years, and especially for those emotion concepts acquired first (happiness, sadness, anger), there is ample evidence to support a role for facial expressions. Nevertheless, the question must be raised concerning exactly how these young children interpret those facial expressions and to what use they put their interpretations. We suggest that before their third birthday, children interpret facial expressions in terms of broad categories, based largely on pleasantness-unpleasantness and, secondarily, arousal (Widen & Russell, 2003a). Facial expressions so interpreted provide the toddler with a broad but relatively undifferentiated structure for emotion.

We thus replicated the Label Superiority Effect: Performance was higher overall given an emotion's label than given its facial expression. The Label Superiority Effect has thus now been found with three different methods (free labeling, sorting, and specifying a cause), suggesting that labels are powerful cues to emotion in different contexts. In the current study, the Label Superiority Effect was significant overall, and strongest for fear and disgust; but weak for sadness; and absent for anger and surprise. These differences across emotions also replicate prior findings (Camras & Allison, 1985; Russell, 1990; Russell & Widen, 2002b). What was especially interesting in the current study was that the effect of labels was significant for 3s and even stronger for 4s. Indeed, improvement with age was limited to the label mode of presentation. Labels might play a more important role in the development of emotion concepts, but mainly when 4-year-olds are acquiring the concepts of fear and disgust. In contrast, at a younger age, and for concepts of happiness,

sadness, and anger, which are acquired earlier (Widen & Russell, 2003a,b), labels were not as powerful—suggesting that labels might play a lesser role. In later development, children acquire concepts for emotions with no strong link to any unique facial expression—such as love, hate, pride, guilt, shame, and so on. Perhaps in these cases, labels will prove to be especially important.

We also found a Behavior Superiority Effect: Performance was higher given an emotion's behavioral consequence than given its facial expression (Fig. 1). This effect was strongest for anger, but present for fear and disgust as well. It was weak for sadness; and reversed for surprise (both label and facial expression were better than the specific consequence we offered children). Our study may have underestimated the power of behavioral consequences. The emotion labels (*sad*, *scared*, etc.) that we presented to the children are likely close to the best instances possible for the label mode of presentation. Similarly, the facial expressions we presented are likely close to the best instances possible, being clear (perhaps even exaggerated) prototypical signals based on Ekman and Friesen's (1978) specifications. The specific labels and facial expressions we used were chosen on the basis of extensive prior work on what is most likely to be understood by this age group. Furthermore, labels presented orally and facial expressions presented visually likely mirror the modality under which children encounter labels and facial expressions in their everyday experience. In contrast, the behavioral consequences we used, although our best effort, were a first effort and might well be subject to improvement through trial-and-error. In addition, the behavioral consequences were presented orally, a modality that although not unprecedented might be less common and less salient for children than witnessing actual behaviors. It may therefore be possible to strengthen the Behavior Superiority Effect by fine-tuning the behavioral consequences used for each emotion and by presenting them pictorially.

Despite these limitations, behavioral consequences were the best of the three modes of presentation overall for 3-year-olds. Behavioral consequences were also especially powerful for the emotion of anger, and anger is one of the earlier emotion concepts acquired, preceded only by happiness (used here as a screening trial) and, in some children, sadness (Widen & Russell, 2003a). The present results offer an intriguing possibility: Especially for 3-year-olds, and especially for anger, an emotion's behavioral consequence might be the most salient cue to forming the concept of that emotion. The present results are the first demonstration of which we are aware that behavioral consequences can be a powerful cue to the emotion of another for a preschooler. Should this effect prove replicable in different contexts and with different methods, it could provide another insight into the younger preschooler's development of emotion concepts.

Taken together, our results suggest a much more interesting developmental story. These results are more in keeping with a script theory than with notions of a single origin for emotion categories. According to a script theory, the features of an emotion (cause, physiology, behavior, facial expression, subjective experience, label, etc.) tend to be correlated; with time and experience, children learn these associations and develop scripts of prototypical emotional events (Bullock & Russell, 1986; Fehr & Russell, 1984). Let us speculate that scripts for happiness and sadness might begin with facial expressions—smiling and crying—that become linked to their causes. For anger, the script might begin instead with hostile behavior being linked to its cause. For fear and disgust, the script might begin with

someone using the labels *scared* and *disgusted*, which becomes linked to their causes. In each case, once the first link is formed the child can then add additional information. Because different cues initiate the process for different emotions, different cues are more powerful for different emotions. Because concepts for different emotions are acquired at different ages, different cues are more powerful at different ages.

The present study was limited to the particular method used, and replication with other methods is required before these hypotheses can be given much credence. For example, with our method, the child told of a causal event for a given cue. It is possible that the child simply associated the causal event directly with the cue, bypassing the emotion altogether. Thus, the child who responded, "someone knocked Danny's tower down" might have imagined this as the cause of the frown (the facial expression of anger), of the use of the word *angry*, or of Danny's reaction of hitting and yelling—all without really understanding that Danny was angry. This possibility can be approached as a methodological problem, and from the network of other research on 3- and 4-year-old children, we are confident that they can be shown to do more than simply associate the causal event with the cue. But there is also a conceptual issue involved. On the script hypothesis just described, a child begins with a pair of events (such as a frustrating event and frowning). From this seed, the script for anger is built by adding more elements. The adult knows anger's subjective feelings and thoughts, its range of behaviors, its physiological manifestations, its social norms, its many contexts, its various means of expression, and so on. More generally, concepts for specific emotions consist of multiple components, which are likely added to the child's knowledge base one at a time. The conceptual question is this: At what point, in this progress from a pair, then a triplet, on to finally a multicomponent adult script does the child "have" the concept of anger? There may be no empirical answer to this question, but rather a fuzzy boundary from not having to having the concept.

Another methodological and conceptual limitation of our study is its focus on children's understanding of the emotions of others. Their understanding of their own emotions was therefore invisible. For example, when witnessing his dad getting angry, Dad's frown may be salient. In contrast, when Chris is angry himself, his own frown (especially its visual appearance) is less salient, perhaps inaccessible. The relative weakness found here for facial expressions is consistent with the possibility that self-experience plays an important role in the development of emotion concepts. Indeed, half of children's spontaneous references to emotions in conversation are to their own emotions (Wellman, Harris, Banerjee, & Sinclair, 1995). This important but difficult-to-study component must also be considered.

Acknowledgments

We thank everyone at the daycares that participated in this study: without their help, this research could not be done. We thank Anita Christy, Aimee Brooks, Diana Garrido, and Stephanie Hadley for their help with data collection and coding. We thank Paul Harris for an extended and enjoyable discussion of the issues raised in this article. We also thank Linda Camras for providing the photographs of facial expressions used in this study. This study was funded by a grant from Boston College.

References

Bandura, A. (1992). Social cognitive theory. In R. Vasta (Ed.), *Six theories of child development: Revised formulations and current issues* (pp. 1–60). London, UK: Jessica Kingsley Publishers.

Barrera, M. E., & Maurer, D. (1981). The perception of facial expressions by the three-month-old. *Child Development, 52*, 203–206.

Barrett, L. F., & Salovey, P. (Eds.). (2002). *The wisdom in feeling: Psychological processes in emotional intelligence*. New York, NY: Guilford Press.

Bowlby, J. (1969). *Attachment and loss: Vol. 1. Attachment* (2nd ed.). London: Hogarth Press.

Bowlby, J. (1988). *A secure base: Parent–child attachment and healthy human development*. New York, NY: Basicbooks.

Bullock, M., & Russell, J. A. (1985). Further evidence on preschooler's interpretations of facial expressions of emotion. *International Journal of Behavioral Development, 8*, 15–38.

Bullock, M., & Russell, J. A. (1986). Concepts of emotion in developmental psychology. In C. E. Izard & P. B. Read (Eds.), *Measuring emotions in infants and children* (Vol. II, pp. 203–237). Cambridge: Cambridge University Press.

Call, J., & Tomasello, M. (1995). Use of social information in the problem solving of orangutans (*Pongo pygmaeus*) and human children (*Homo sapiens*). *Journal of Comparative Psychology, 109*, 308–320.

Camras, L. A., & Allison, K. (1985). Children's understanding of emotional facial expressions and verbal labels. *Journal of Nonverbal Behavior, 9*, 84–94.

Camras, L. A., Grow, J., & Ribordy, S. (1983). Recognition of emotional expression by abused children. *Journal of Clinical Child Psychology, 12*, 325–328.

Caron, A. J., Caron, R. F., & MacLean, D. J. (1988). Infant discrimination of naturalistic emotional expressions: The role of face and voice. *Child Development, 59*, 604–616.

Darwin, C. (1965). *The expression of the emotions in man and animals*. Chicago: University of Chicago Press. Original work published in 1872.

De Rosnay, M., & Harris, P. L. (2002). Individual differences in children's understanding of emotion: The roles of attachment and language. *Attachment and Human Development, 4*, 39–54.

Denham, S. A. (1998). *Emotional development in young children*. New York: Guilford Press.

Denham, S. A., & Couchoud, E. A. (1990). Young preschoolers' ability to identify emotions in equivocal situations. *Child Study Journal, 20*, 153–169.

Ekman, P., & Friesen, W. V. (1978). *Facial action coding systems*. Palo Alto, CA: Consulting Psychologists Press.

Fehr, B., & Russell, J. A. (1984). Concept of emotion viewed from a prototype perspective. *Journal of Experimental Psychology: General, 113*, 464–486.

Gnepp, J. (1983). Children's social sensitivity: Inferring emotions from conflicting cues. *Developmental Psychology, 19*, 805–814.

Harré, R. (1986). *The social construction of emotions*. New York, NY: Blackwell.

Harrigan, J. A. (1984). The effect of task order on children's identification of facial expressions. *Motivation and Emotion, 8*, 157–169.

Harris, P. L. (1989). *Children and emotion: The development of psychological understanding*. Oxford, UK: Basil Blackwell.

Heyman, G. D., & Gelman, S. A. (1998). Young children use motive information to make trait inferences. *Developmental Psychology, 34*, 310–321.

Izard, C. E. (1971). *The face of emotion*. New York: Appleton Century Crofts.

Izard, C. E. (1994). Innate and universal facial expressions: Evidence from developmental and cross-cultural research. *Psychological Bulletin, 2*, 288–299.

Klinnert, M. D., Emde, R. N., Butterfield, P., & Campos, J. J. (1988). Social referencing: The infant's use of emotional signals from a friendly adult with mother present. In S. Chess, A. Thomas, & M. Hertzig (Eds.), *Annual progress in child psychiatry and child development* (pp. 26–39). New York: Brunner/Mazel.

Lutz, C. (1988). *Unnatural emotions: Everyday sentiments on a Micronesian atoll & their challenge to western theory*. Chicago, IL: University of Chicago Press.

Magai, C., & McFadden, S. H. (1995). *The role of emotions in social and personality development: History, theory, and research*. New York: Plenum Press.

Markham, R., & Adams, K. (1992). The effect of type of task on children's identification of facial expressions. *Journal of Nonverbal Behavior, 16*, 21–39.

Maurer, D., & Barrera, M. E. (1981). Infants' perception of natural and distorted arrangements of a schematic face. *Child Development, 52*, 196–202.

Moses, L. J., Baldwin, D. A., Rosicky, J. G., & Tidball, G. (2001). Evidence for referential understanding in the emotions domain at twelve and eighteen months. *Child Development, 72*, 718–735.

Powlishta, K. K., Sen, M. G., Serbin, L. A., Poulin-Dubois, D., & Eichstedt, J. (2001). A from infancy through middle childhood: The role of cognitive and social factors in becoming gendered. In R. K. Unger (Ed.), *Handbook of the psychology of women and gender* (pp. 116–132). New York, NY: Wiley.

Reichenbach, L., & Masters, J. C. (1983). Children's use of expressive and contextual cues in judgements of emotion. *Child Development, 54*, 993–1004.

Robinson, J. L., & Acevedo, M. C. (2001). Infant reactivity and reliance on mother during emotion challenges: Prediction of cognition and language skills in a low-income sample. *Child Development, 72*, 402–415.

Russell, J. A. (1990). The preschooler's understanding of the causes and consequences of emotion. *Child Development, 61*, 1872–1881.

Russell, J. A., & Widen, S. C. (2002a). A Label Superiority Effect in children's categorization of facial expressions. *Social Development, 11*, 30–52.

Russell, J. A., & Widen, S. C. (2002b). Words versus faces in evoking children's knowledge of the causes of emotions. *International Journal of Behavioral Development, 26*, 97–103.

Serrano, J. M., Iglesias, J., & Loeches, A. (1992). Visual discrimination and recognition of facial expressions of anger, fear, and surprise in 4- to 6-month-old infants. *Developmental Psychobiology, 25*, 411–425.

Serrano, J. M., Iglesias, J., & Loeches, A. (1995). Infants' responses to adult static facial expressions. *Infant Behavior and Development, 18*, 477–482.

Shields, A., Dickstein, S., Seifer, R., Giusti, L., Magee, K. D., & Spritz, B. (2001). Emotional competence and early school adjustment: A study of preschoolers at risk. *Early Education and Development, 12*, 73–96.

Smith, M. (2001). Social and emotional competencies: Contributions to young African-American children's peer acceptance. *Early Education and Development, 12*, 49–72.

Surbey, P. D. (1979). *Preschool children's understanding of emotional states in terms of causes and consequences*. Unpublished master's thesis, University of Minnesota, cited in Russell (1990).

Walker-Andrews, A. S., & Lennon, E. (1991). Infants' discrimination of vocal expressions: Contributions of auditory and visual information. *Infant Behavior and Development, 14*, 131–142.

Want, S. C., & Harris, P. L. (2001). Learning from other people's mistakes: Causal understanding in learning to use a tool. *Child Development, 72*, 431–443.

Wellman, H. M., Harris, P. L., Banerjee, M., & Sinclair, A. (1995). Early understanding of emotion: Evidence from natural language. *Cognition and Emotion, 9*, 117–149.

Widen, S. C., & Russell, J. A. (2002). Gender and preschoolers' perception of emotion. *Merrill-Palmer Quarterly, 48*, 248–262.

Widen, S. C., & Russell, J. A. (2003a). A closer look at preschoolers' freely produced labels for facial expressions. *Developmental Psychology, 39*, 114–128.

Widen, S. C., & Russell, J. A. (2003b). Free labeling of emotions by two- and three-year-olds: A face superiority effect. Submitted for publication.

Wierzbicka, A. (1992). Defining emotion concepts. *Cognitive Science, 16*, 539–581.

Wiggers, M., & van Lieshout, C. F. M. (1985). Development of recognition of emotions: Children's reliance on situational and facial expressive cues. *Developmental Psychology, 21*, 338–349.