Available online at www.sciencedirect.com
ScienceDirect
Cognitive Development 23 (2008) 291–312
COGNITIVE DEVELOPMENT

Children acquire emotion categories gradually

Sherri C. Widen *, James A. Russell

Boston College, Chestnut Hill, MA, USA

Abstract

Some accounts imply that basic-level emotion categories are acquired early and quickly, whereas others imply that they are acquired later and more gradually. Our study examined this question for *fear, happiness, sadness*, and *anger* in the context of children's categorization of emotional facial expressions. Children ($N = 168$, 2–5 years) first labeled facial expressions of six emotions and were then shown a box and asked to put all and only, e.g., scared people in it. Before using *fear* in labeling, children had begun to include 'fear' faces and to exclude other (especially positive) faces from the fear box/category; after using *fear*, children continued to include other (especially negative) faces. The same pattern was observed for happiness, sadness, and anger. Emotion categories begin broad, including all emotions/faces of the same valence, and then gradually narrow over the preschool years.

Keywords: Categories; Labeling; Emotion; Preschoolers

Childhood presents a number of landmarks, such as the child's first word or first use of a particular word. Landmarks are then used to chart development—as when a child's first use of the word *scared* is thought to mark the acquisition of the concept of fear. Even if they do not mean to, when psychologists write of children 'acquiring' a word or concept based on a landmark event, they seem to imply an all-or-none achievement. In this article, we explore one set of such landmarks: children's acquisition and use of emotion labels. Although use of a label is a discrete event (the child either does or does not use, for example, the word *scared*), acquisition of an emotion concept may be quite gradual. This hypothesis stems from our previous work suggesting that emotion concepts are initially broad, including anything of the same valence, and narrow gradually over a period of years (Russell & Paris, 1994; Widen & Russell, 2003).

Evidence regarding children's acquisition of emotion concepts presents a puzzle. On the one hand, many believe that very young infants already recognize and respond to different discrete

* Corresponding author at: Department of Psychology, McGuinn Hall, 140 Commonwealth Avenue, Boston College, Chestnut Hill, MA 02467, USA. Tel.: +1 617 552 6768; fax: +1 617 552 0523.
E-mail address: widensh@bc.edu (S.C. Widen).

0885-2014/$ – see front matter © 2008 Elsevier Inc. All rights reserved.
doi:10.1016/j.cogdev.2008.01.002

emotions perceived via facial and vocal expressions (Denham, 1998; Izard, 1971, 1994; Walker-Andrews & Lennon, 1991). These babies are thought to recognize when their caregiver is happy, frightened, or sad, for example. Darwin and those who followed in his footsteps suggested that the set of categories for at least basic-level emotions are a universal heritage from our evolutionary past. On this view, the child needs simply to find the right label for a category already possessed. The 'fast mapping' notion in acquisition of word meaning implies that even if not all-or-none, acquisition is fast (e.g., Dollaghan, 1985; Heibeck & Markman, 1987; Rice, 1990).

Consistent with the idea of early and fast acquisition, Dunn, Bretherton, and Munn (1987) concluded that 2-year olds could "distinguish and discuss" (p. 139) a variety of emotions. Indeed, children begin using emotion labels before their second birthday (e.g., Ridgeway, Waters, & Kuczaj, 1985), and two-year olds' use of these words suggests a surprisingly sophisticated understanding of emotion (Wellman, Harris, Banerjee, & Sinclair, 1995). Observational studies show that even two-year olds talk about their own and others' emotions (Bretherton, Fritz, Zahn-Waxler, & Ridgeway, 1986; Wellman et al., 1995), attribute emotions to dolls and toy animals (Wolf, Rygh, & Altshuler, 1984) and to their siblings during pretend play (Dunn et al., 1987), and use emotion language both to tease and to manipulate others' behaviors (Bretherton, McNew, & Beeghly-Smith, 1981; Wellman et al., 1995). Wellman et al. (1995) concluded that children understand that emotions are internal feelings, distinct from the causes that elicit such feelings and from the behaviors and expressions that result from them. Altogether, such theory and evidence have been taken to support the assumption of an early and sophisticated understanding of at least basic-level emotions, especially when given biologically based cues such as facial expressions. The implication is that basic-level emotion categories are in place early, at least by the third year, especially when the child is given facial expressions as the cue to emotion.

On the other hand, when children's emotion categories are scrutinized in more structured tasks, even older preschoolers' concepts appear less complete than might be expected. Consider, for example, studies that have compared children's performance on a number of different emotion tasks: free labeling, forced choice from an array of labels, matching facial expressions from an array, and matching a facial expression to a brief story that included the emotion label (Markham & Adams, 1992; Vicari, Reilly, Pasqualetti, Vizzotto, & Caltagirone, 2000). Four- and five-year olds performed as well as grade-schoolers for happiness, but their performance was significantly lower, in descending order, for sadness, anger, fear, surprise, and disgust. Markham and Adams proposed that four-year olds' lower performance for these five emotions was not due to simple performance difficulties, but rather, to the fact that their emotion categories were still developing. Vicari et al. concluded that recognition of facial expressions continues to develop well into the school years.

Emotion categories are not unique in raising the question of whether and how well children understand these categories even after using the associated labels. Similar questions have also been raised in the area of dimensional adjectives for color and size. For example, although children begin using color adjectives before their 3rd birthday, they use them accurately only at 4 or 5 years of age (Bornstein, 1985). Longitudinal research has shown that an adult-like understanding of size and color adjectives requires repeated exposure and many opportunities for learning (e.g., Carey, 1985; Carey & Bartlett, 1978; Roberson, Davidoff, Davies, & Shapiro, 2004; Sandhofer & Smith, 1999).

A gradual emergence of categories is consistent with, and perhaps at least in part explained by, another finding: 'errors' are systematic rather than random. Analysis of errors suggests underlying continuous dimensions. Children's categorization 'errors' on emotion tasks can be predicted by the similarity of the stimuli in terms of their general pleasant versus unpleasant qualities (Bullock

& Russell, 1984). Likewise, in Roberson et al.'s (2004) study on children's acquisition of color concepts, children's errors on a memory task were predicted by perceptual distance. Thus, even after the children began using the relevant labels, their understanding of the associated emotion or color concepts continued to develop for a considerable period of time; and errors were based on the similarity of stimuli along underlying dimensions.

In summary, there are conflicting assumptions about the timing of children's acquisition of emotion concepts (just as for concepts of color and size). Our concern is not simply timing, of course, but the nature of the acquisition process, the nature of the constraints on acquisition, and the actual concepts children use during an important time in their development. From our perspective, children do not understand emotions, their own and those they witness, in terms of adult-like discrete-categories they acquire early and quickly—as commonly assumed in the literature. Instead, over the preschool years, children move gradually from understanding emotions in very broad terms to narrower ones, thus slowly moving toward an adult-like understanding.

In this article, we explore this issue in the context of a Differentiation Model, which attempts to describe children's acquisition of emotion concepts (Widen & Russell, 2003). In this model, children initially understand emotions in terms of the broad dimensions of pleasure–displeasure and degree of arousal – which underlie the structural circumplex model of emotion (Russell, 1980; Fig. 1); similarity on these two dimensions allows emotions to be arrayed in a structure, which turns out to be a circle in the two-dimensional geometric pleasure–arousal space – and these dimensions remain a major part of emotion perception in adults. Early on, children form categories of emotion, initially based on these broad dimensions but later by incorporating other information.

One step in developing the Differentiation Model was charting the order in which emotion concepts are acquired. Originally, we thus assumed that the first observed use of an emotion label in free labeling marked the child's acquisition of the corresponding emotion category. Here, we explore the idea at a deeper level: Although the first observed use of an emotion term is a discrete event, and an important one, the underlying process is a gradual narrowing of the concept.

Initial data supporting the Differentiation Model were children's freely produced labels, both 'correct' and 'incorrect,' for prototypical facial expressions of emotion (Widen & Russell, 2003).

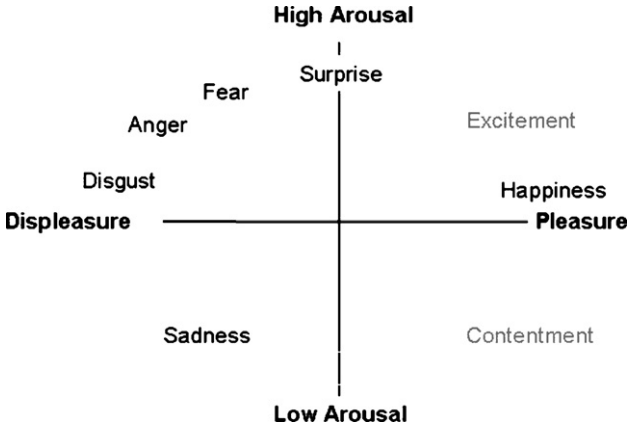

Fig. 1. The circumplex model of emotion, including the six emotions used in both the free labeling and categorization tasks (printed in black), and the additional three emotions used only in the categorization task (printed in grey). Adapted from Russell (1980).

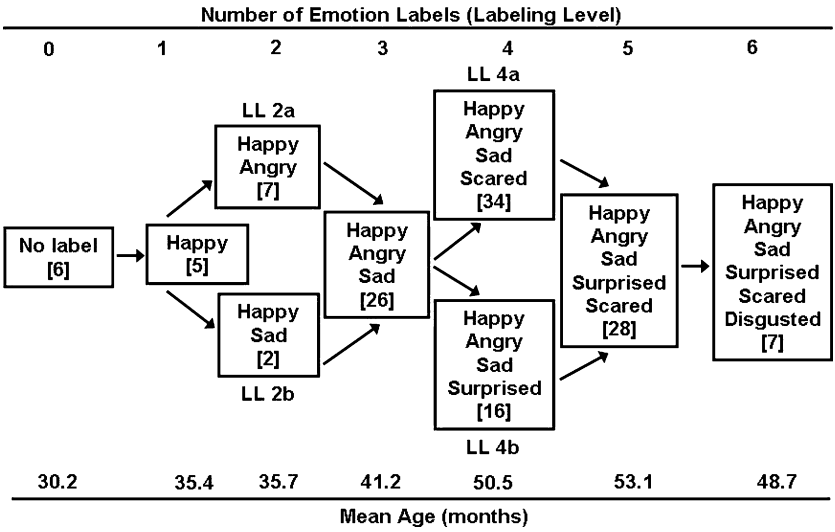

Fig. 2. Systematic emergence of emotion labels. The number of children who used the specified set of labels is given in brackets [*n*].

These data reflect four specific findings. First, from 2 to 5 years, children's 'errors' are systematic rather than random: Children were more likely to 'mislabel' a face with a label from a similar emotion category than from a dissimilar one, with similarity specified by our structural circumplex model (Fig. 1) (Widen & Russell, 2003).

Second, when children (2–5 years of age) were classified, irrespective of age, by the number of emotion category labels they used, emotion labels were found to emerge in a systematic order, as illustrated in Fig. 2 (Widen & Russell, 2003). Over 81% of the children fit this pattern. Age increased with the number of labels used from a mean age of 30 months for Labeling Level 0 (children who produced no labels) to 62 months for Labeling Level 5 (children who used five labels).

Third, children used different emotion labels with different frequencies, even when presented with an equal number of stimuli (facial expressions or stories including causes and consequences) for each emotion (Nelson, Widen, & Russell, 2006; Widen & Russell, 2003). Labels for earlier-emerging categories (happiness, sadness, anger) were used most frequently, labels for later-emerging categories (fear, surprise, disgust) less frequently. We also found a similar order for 'incorrect' labels (Widen & Russell, 2003, Study 2). Differential use of labels was therefore not a result of the faces per se, but rather of children's category system for emotion. We interpreted this finding as an indication that early-emerging categories remain broader than later-emerging ones.

Fourth, emotion categories narrow with age (Widen & Russell, 2003). Two- and three-year olds used emotion category labels for the target faces *and* for a wide variety of other faces as well—they used *happiness* for happiness, surprise, and fear faces (Widen & Russell, 2003). Four- and five-year olds used emotion labels for the target face, but for a narrower range of other faces.

The Differentiation Model is now simply descriptive, based, as it is, on a single task that requires production of labels: free labeling of facial expressions. In the current study, we replicate and extend the model with data from a categorization task requiring only comprehension of emotion labels. We hope to show that the comprehension task provides similar evidence supporting the Differentiation Model, revealing children's underlying emotion categories.

But more importantly, we extend the model. To our knowledge, the relationship between the acquisition of an emotion label and the development of a child's emotion categories has not previously been systematically investigated. One possibility is that when children begin to use *fear*, for example, in free labeling they will show a sharp shift in categorization performance, revealing their greater understanding of fear relative to children at lower Labeling Levels. There is evidence that labels are powerful cues to emotion (Russell & Widen, 2002a), especially for older preschoolers and for later-emerging emotion categories such as fear and disgust (Camras & Allison, 1985; Russell, 1990; Widen & Russell, 2004). The second possibility is that children's categorization will narrow gradually both before and after they use *fear* in free labeling, because label use is only one part of the acquisition of the category of fear. The key question for our purposes is: Does the first observed use of a label in the free labeling task mark a qualitative change in the nature of the underlying category?

1. Method

1.1. Participants

Participants were 168 children, all proficient in English and enrolled in 25 preschools in the Greater Boston, MA, area. There were 42 boys and 42 girls in each of two age groups: Young (24–47 months; mean = 38 months, S.D. = 5.8; 33% of the children were 24–36 months; 67% were 36–47 months) and older preschoolers (48–65 months; mean = 54 months, S.D. = 4.0; 59% of the children were 27–55 months; 41% were 56–65 months). In total, the children were 81% European-American, 8% mixed-race, 6% Asian-American, and 5% other ethnicities. On the consent form, parents were asked indicate the highest level of education completed by each parent (to assess SES) on a six-point scale: (1) some high school; (2) high school diploma or GED; (3) some college, vocational degree, associates degree; (4) 4-year college degree; (5) master's degree; and (6) post-master's degree (PhD, MD, MBA, JD, EdD, ThD). Although education alone is not a sufficient indicator of socioeconomic status, it has been used in the past as a major component of indices of SES (Hollingshead & Redlich, 1958). Parents' mean education level for this sample was 4.9 (S.D. = 1.1). A sample of 24 university-aged adults were also included; they received course credit in exchange for their participation.

1.2. Design

Each child participated in the three parts of the study in the same order: free labeling animals, free labeling faces, categorization of faces. First, children free labeled photographs of three animals; this phase served as a training trial and a comparison task for free labeling facial expressions; it established that the children were willing and able to produce labels on demand. Second, children free labeled two sets of six facial expressions ('happiness,' 'sadness,' 'anger,' 'fear,' 'surprise,' 'disgust'). This phase provided a measure of which emotion labels these children produced spontaneously. The number of different emotion category labels that each child used established his or her Labeling Level. In the free labeling procedure, the experimenter shows a child a still photograph of a facial expression and asks, "How is this person feeling?" One advantage of free labeling is that it comes closest to tapping a child's spontaneous specification of the emotion seen in a face.

Third, in the categorization task, children categorized each of eight faces into or out of the target emotion category for each trial. An emotion category was presented to the child as a physical

analogue: a box. The target emotion was verbally specified by a label (*happy*, *sad*, *angry*, or *scared*). (Nevertheless, evidence shows that the box task yields similar results when the box is defined nonverbally by showing facial expressions of the corresponding emotion; Russell & Widen, 2002b.) The child was told that people who feel only a certain way (e.g., *scared*) could go in the box. The child's task was to decide whether each of eight facial expressions (test faces) should go into the box or be left out. The test faces were 'surprise,' 'excitement,' 'happiness,' 'contentment,' 'sadness,' 'disgust,' 'anger,' and 'fear.' Three positive facial expressions were included on each trial for two reasons. The first is to round out the circumplex (Fig. 1). Without the 'excitement' and 'contentment' faces, the negative emotions in the circumplex (fear, anger, disgust, sadness) are given undue weight. Second, piloting indicated that, on the sadness, anger, and fear trials, most children included other negative expressions. Thus, including more than one positive expression made the happiness trial more similar to the other trials. Children were not told prior to the categorization task how many trials there were nor which specific target categories would be tested. We did not include three positive facial expressions in the free labeling task, which would have made that task more parallel to the categorization task, because doing so would have given undue emphasis to happiness. On the free labeling task, children label each face in turn, and most will use the label *happiness*. Thus, having three positive facial expressions in each set of faces would have offered no advantage, and might have created a labeling bias (some children might have learned to label any face they were unsure of as *happiness* because there were so many happiness exemplars) or a reluctance to label later happiness faces *happiness* (because there was only one of each of the other expressions).

1.3. Materials

1.3.1. Photographs of animals

The animal pictures were three color photographs, one each of a cat, dog, and rabbit.

1.3.2. Photographs of facial expressions for free labeling

Two sets each of seven black-and-white 5 in. × 7 in. photographs (one set posed by a boy, one by a girl) of prototypical facial expressions of emotion ('happiness,' 'sadness,' 'anger,' 'fear,' 'surprise,' 'disgust,' 'neutral') were used in free labeling.

1.3.3. Photographs of facial expressions for categorization

Four sets each of eight facial black-and-white 5 in. × 7 in. photographs of prototypical facial expressions ('surprise,' 'excitement,' 'happiness,' 'contentment,' 'sadness', 'disgust,' 'anger,' 'fear'), each posed by an adult woman, were used in the categorization task.[1] None of the faces

[1] The photographs used in the free labeling facial expressions task were provided by Dr. Linda Camras. Camras et al. (1983) described the development of the photographs, their coding according to Ekman and Friesen's (1978) facial action coding system (FACS), and their use in a study on recognition of emotional expressions. Twenty-three of the facial expressions used in the categorization task were selected from Ekman and Friesen's (1976) pictures of facial affect; four from Matsumoto and Ekman's (1988) JACFEE collection; and five from our own collection. In the free labeling facial expressions task, we used photographs posed by children because they have been FACS coded and were used in prior studies. In addition, for each model there was a complete set of facial expressions (six basic-level expressions, plus neutral), which in free labeling allows the child to focus on the facial expression rather than differences in models' age, race, and physical appearance. In the categorization task, we used photographs posed by adult women, again because they have been FACS coded and were used in prior studies. For the Categorization task, we needed 32 different photographs,

used in the categorization task were used in free labeling. No facial expression was used twice in the categorization task.

1.4. Procedure

The experimenter spent the first visit getting to know each child. On a subsequent visit, the experimenter invited an individual child to play a game with her.

For purposes unrelated to the present work, the experimenter first engaged the child in a brief conversation about emotions in which each of the target emotion labels were introduced or elicited (one third of the children participated in a conversation about colors). For a description of these 'priming' procedures, see Widen and Russell (2003, Studies 1 and 3). Comparison of children who experienced priming procedures and those who did not revealed no significant differences in subsequent performance.

1.4.1. Animal labeling

The animal labeling trials served as a practice session and a comparison task for the free labeling of facial expressions. The experimenter said, "Do you know what I brought with me today? I brought some pictures of animals. Would you like to see them?" The three animals were shown, in different random orders. For each one, the experimenter asked, "Do you know what kind of animal this is?" Responses were not corrected and were mildly praised equally (e.g., "Good answer;" "You are very good at this game"). The experimenter ended by saying, "That was really fun. Do you know what else I brought?"

1.4.2. Free labeling facial expressions

The order of presentation for the two sets of facial expressions was counterbalanced: Half the children saw Jack first; half Sally. (Each group was evenly divided by gender.) The experimenter introduced the faces by saying, "I brought some pictures of a boy named Jack (a girl named Sally). Do you want to see them? Okay, here is a picture of Jack (Sally) (showing the neutral expression). Do you know what Jack (Sally) is going to do? He (she) is going to show us how he (she) feels sometimes. Let's see if you can tell me how Jack (Sally) feels in each one." The experimenter then showed the child the six emotional facial expressions, one at a time in a random order. For the first face, the experimenter said, "One day, Jack (Sally) felt like this (pointing to the face)." For the other faces, the experimenter said, "One week later, Jack (Sally) felt like this (pointing to the picture)." After each picture, the experimenter asked, "How do you think Jack (Sally) feels in this picture?" Responses were not corrected and all were mildly praised (e.g., "Good answer"; "You are good at this game."). If no response was given, the experimenter used various prompts (Have you ever made this face? What do you think happened to make Sally feel this way?). If the child still did not respond, the experimenter went on to the next photograph, and, after the other trials for both sets of photographs, returned to any to which the child had not responded. At no time did the experimenter use the word *emotion*, provide any other emotion label, or otherwise direct the child to try to use an emotion label beyond asking how Jack (Sally) was feeling. After seeing all six of the first set of faces, the experimenter introduced the second set: "That was great.

---

and those posed by adults are more plentiful. Piloting with these photographs indicated that the children performed well. All the photographs in both tasks were clear, prototypical facial expressions of basic-level emotions. There is no evidence that photographs posed by children vs. adults yield different results.

Do you know what else I brought with me today? I brought some pictures of a girl named Sally (boy named Jack). Would you like to see them?"

1.4.3. Categorization task

Each child participated in four categorization trials, one each for happiness, sadness, anger, and fear. The trials were presented in different random orders. In a prior study, children as young as 2 years completed a control categorization trial (color) prior to the emotion trials (Russell & Widen, 2002b). In that study, only 11 of 360 (3.1%) children made any errors on the control task, and so we were confident that the children in the current study could do this task. The category was defined as a box into which only people who felt a certain way could go. On each trial, the child was shown a different set of eight facial expressions: one each of 'excitement,' 'happiness,' 'contentment,' 'sadness,' 'disgust,' 'anger,' 'fear,' and 'surprise,' presented one at a time in random order. For example, on the happiness trial, experimenter explained, "This is a special box. It is only for happy people, and only happy people can go in the box. All the other people go out here (pointing to a spot beside the box)." The experimenter then showed the child the eight test photographs, one at a time in random order. For each test expression, the experimenter asked, "Is this person happy? Does she go in the happy box or out?" If the child did not answer both questions consistently (e.g., that a fear expression was *not* happy and should *not* go in the box), he or she was reminded of the rules, and the questions were repeated. 'Incorrect' categorization was not corrected; all answers were mildly praised. A new box was used for each trial.

1.4.4. Adult comparison group

Adults completed the free labeling facial expressions and categorization tasks in a questionnaire format. Adults were included as a comparison to establish developmental endpoints for the two tasks. Free labeling facial expression was presented first. For half the adults, Jack was first, for half, Sally; each set of faces was presented in a different random order. Participants were asked to label the emotion in each facial expression, with one word if possible. The categorization task followed. For the adults, all eight faces for each trial were displayed in an array of two rows. There was a circle above each face in the top row and below each face in the bottom row. On each trial (happy, sad, angry, scared), participants were asked to put an X in the circle(s) for the face(s) that displayed the target emotion. The faces in free labeling and the categorization trials were presented in different random orders.

1.5. Scoring

1.5.1. Animals

The labels scored as correct in the cat category were *cat, kitten*; in the dog category, *dog*, *doggy, puppy*; in the rabbit category, *rabbit*, *bunny*. Children used no other labels.

1.5.2. Free labeling facial expressions

On the free labeling task, participants were allowed to use any emotion label they chose. The scoring key used in this study was drawn from Widen and Russell (2003), who describe the development of a scoring key based on ratings of two judges blind to the source of the labels. The labels that children used that were scored as correct for each category were: for happiness, *happy, good*, *exciting*, *great*; for fear, *scared, frightened*; for disgust, *disgusted*, *yucky*, *ew*, *gross*, *icky*; for anger, *angry*, *mad*, *grumpy*, *frustrated*; for surprise, *surprised*, *shocked*; and for sadness, *sad*, *upset*, *depressed*. Responses varied in syntax or by being embedded in a phrase (e.g., *very*

scared). Responses that did not fit into one of the target emotion categories were: *a ghost*; *awful*; *bad*; *bored*; *cold*; *cranky*; *crazy*; *crying*; *dizzy*; *don't know*; *eat food*; *fat*; *feels dear*; *felt like her mom put sauce on her food*; *freezing*; *funny*; *goofy*; *he's a monster*; *her feels like eyes*; *hot*; *kinda coughing*; *laugh*; *making her teeth open*; *mean*; *not feel so well*; *not happiness*; *not nice*; *oh my gosh*; *open her mouth*; *play ball*; *quiet*; *scary*; *screaming*; *serious*; *shouting*; *silly*; *smelly*; *smiling*; *stink*; *stumped*; *sun's in her eyes*; *sunny*; *terrible*; *tired*; *weird*; *yawned*. These were all the labels children used in the current study.

1.5.3. Categorization task

Responses on the categorization task were scored in two different ways. In the first scoring method, responses were scored as correct if the target face was included in the category (e.g., placing the anger face in the anger box; for adults, marking the anger face with an 'x') and if any nontarget faces were excluded from the category (e.g., placing any of the 'surprise,' 'excitement,' 'happiness,' 'contentment,' 'sadness,' 'disgust,' and 'fear' faces outside of the anger box; for adults, not marking any of these faces), and incorrect if the target face was excluded from the box (e.g., placing the 'anger' face outside of the anger box; for adults, not marking the anger face) or if any nontarget faces were included in the category (e.g., placing any of the 'surprise,' 'excitement,' 'happiness,' 'contentment,' 'sadness,' 'disgust,' and 'fear' faces in the anger box; for adults, marking any of these faces). For the happiness box, inclusions of the 'excitement,' 'happiness,' and 'contentment,' faces were scored as correct, as each of these faces displayed varying levels of happiness.

In the second scoring method, responses were scored as inclusions and exclusions (1 and 0, respectively) in order to measure the breadth of children's emotion categories. Thus, any faces that were placed in a box were scored as inclusions, and those that were left out of the box were scored as exclusions. For adults, inclusions were all faces that had been marked with an 'x'; those that were not marked with an 'x' were exclusions.

2. Results

2.1. Free labeling task: animal and facial expression labeling

Every child labeled every animal correctly, showing that all the children in our study could and did free label. Collectively, the 168 children had 2016 opportunities to label a facial expression. Overall, the children did so 'correctly' on 1075 of those opportunities. Most children performed moderately: 70.2% gave the target response for 4–8 of the 12 faces. Proportion of correct emotion labels (53.7%) was lower than proportion (100%) of correct animal labels, and also lower than the proportion of correct emotion labels (90.6%) produced by adults for the same faces. These results are within the range of previous free labeling results for preschoolers (e.g., Harrigan, 1984; Markham & Adams, 1992; Widen & Russell, 2002, 2003).

2.2. Feeling as a lexical class

To our knowledge, no data have previously been reported on the question of whether *feeling* is a lexical class for preschoolers. A stringent test of children's knowledge of a lexical class of feeling labels is provided by considering only their 'incorrect' responses, since all correct responses are emotion labels by definition. Children responded incorrectly on 941 trials. On these trials, children provided an emotion label on 644 (68.4%) trials, a non-emotion feeling word on 66 (7.0%) trials, a

	Response Category						
Face	Happiness	Surprise	Fear	Anger	Disgust	Sadness	Total
'Happiness'	**302**	1	1	3	0	2	309
'Surprise'	39	**160**	41	13	2	9	264
'Fear'	63	33	**91**	33	11	44	275
'Anger'	5	1	5	**264**	4	16	295
'Disgust'	8	3	5	192	**30**	34	272
'Sadness'	7	2	3	64	0	**228**	304
Total	424	200	146	569	47	333	1719
Errors	122	40	55	305	17	105	644
(%)	(28.8)	(20.0)	(37.7)	(53.6)	(29.8)	(31.5)	(37.5)

Fig. 3. The number of times each face was labeled with a response from each emotion response category. The emotion categories and facial expressions are ordered in accordance with a structural model based on similarity of pleasantness and arousal. All the "correct" responses are on the center diagonal and the "errors" are on the other diagonals. 'Errors' one step removed are on the two diagonals adjacent to the center white one, and so on.

non-feeling response on 114 (12.1%) trials, and a non-response on 117 (12.4%) trials. Thus, even when correct responses were removed from consideration, children's responses to the question of how Sally or Jack is feeling were feeling labels on 75.4% of trials.

2.3. *Four findings relevant to the differentiation model*

2.3.1. *Systematic errors*

We first asked whether children's 'errors' in the free labeling task were systematic. Fig. 3 shows the number of times each label was applied to each facial expression. The prediction was that the likelihood of using a label for a given face is inversely proportional to the number of steps away from the main (white) diagonal. To test this prediction, we summed the cells of each diagonal and divided it by the number of cells in that diagonal. For example, the relative frequency in the Step 1 diagonal = [1 + 41 + 33 + 4 + 34 + 39 + 33 + 5 + 192 + 0]/10 = 382/10 = 38.2; in Step 2, 21.8; in Step 3, 10.0; in Step 4, 4.8; and in Step 5, 4.5. The prediction was confirmed by the decreasing relative frequency as the number of steps increased: Errors were not random, but were more likely to be from emotion categories closer to the target category than from categories further removed.[2]

2.3.2. *Systematic emergence of emotion labels/categories*

To explore whether emotion labels/categories emerged systematically during development, all children were classified, irrespective of age, by the number of different emotion category labels they used. For each number, the frequency with which each possible combination of emotion labels occurred was counted. Labels emerged in the predicted pattern (Fig. 2). Six children used

[2] Because fully 50% of cases in Step 1 were due to the children labeling the 'disgust face' as *anger*, we recalculated the relative frequencies without the 'disgust face' and the disgust response category. Even without this high frequency error in Step 1, the first finding of the Differentiation Model was confirmed, and the relative frequency of errors decreased as Steps away from the center diagonal increased, although Steps 3 and 4 were virtually the same size: Correct = 1045/5 = 209.0; Step 1 = 232/8 = 29.0; Step 2 = 125/6 = 20.8; Step 3 = 19/4 = 4.8; and Step 4 = 9/2 = 4.5.

no labels (Labeling Level 0). If a child used only one label, that label was most likely to be *happiness* (Labeling Level 1); 5 of 7 children (71%) who used only one label used *happiness*. For two labels, 9 of 11 (82%) who used two labels used *happiness* and *sadness* or *happiness* and *anger* (Labeling Level 2); for three labels, 26 of 34 (76%) who used three labels used *happiness*, *sadness*, and *anger* (Labeling Level 3); for four labels, 50 of 62 (83%) who used four labels added either *surprise* or *fear* (Labeling Level 4); for five labels, 28 of 41 (68%) who used five labels used *happiness*, *sadness*, *anger*, *fear*, and *surprise* (Labeling Level 5); and for six labels, seven children added *disgust* (Labeling Level 6). Frequencies for each of the non-predicted combinations were low. The model accounted for 78.0% (131) of the children—a proportion significantly greater ($p < .001$) than the 19.3% expected by chance. (That is, if, as the number of labels that children used increased, any label was as likely to be added as any other, then 19.3% would fit the predicted pattern.) Age increased with Labeling Level from Labeling Level 0 to Labeling Level 5 (Fig. 2, bottom). The mean difference in age between these Labeling Levels was 4.6 months. All but one adult (95.8%) fit the predicted pattern: Conforming were 22 of the 24 adults who used all six target emotion labels plus one who used all but *disgust*; non-conforming was one adult who omitted *surprise*.

2.3.3. Frequency of use of specific labels

Each child was presented with an equal number of facial expressions (two) for an equal number of allegedly discrete emotions (six), and yet children used some labels more frequently than others (Fig. 3, column totals). *Anger* was used most frequently, followed, in rank order, by *happiness*, and *sadness*, and then *surprise*, *fear*, and *disgust*. This general pattern was replicated for both correct responses (on the center diagonal), with a reversal in rank order between *happiness* and *anger*, and 'incorrect' responses, with a reversal in rank order between *fear* and *surprise*.

2.3.4. Narrowing of emotion categories with age

The final finding relevant to the Differentiation Model showed that children's emotion categories begin broad and narrow gradually. In addition, earlier-emerging emotion categories (happiness, sadness, anger) are initially broader than later-emerging ones. Both observations were supported by the free labeling results: The young preschoolers used *happiness* primarily for the happiness faces (81.5% of trials), and also for the surprise (13.7%) and fear (22.6%) faces. The older preschoolers used *happiness* for the 'happiness' faces (98.2%) even more frequently—a significant age difference in the average proportion correct on happiness trials, $t_{166} = 4.00$, $p < .001$. Older preschoolers used *happiness* less frequently for the 'surprise' (9.5%) and 'fear' (14.9%) faces than young preschoolers; although the trend was in the expected direction, it was not significant ($ps > .11$); use of *happiness* was near floor levels for the other faces for both groups. A similar pattern was observed for *sadness* and *anger*: all three of the earliest-emerging categories began exceedingly broad; with age, children's early emotion categories narrowed, becoming more adult-like.

On this production task, children's later-emerging categories (fear, surprise, disgust) did not show the same initial breadth as their earliest emotion categories. For example, young preschoolers used *surprise* on average on only 6.3% of all opportunities to label a face (compared to 21.4% for *happiness*), and they used *surprise* almost exclusively for the 'surprise' (76.2%) and 'fear' (17.5%) faces. Older preschoolers used *surprise* significantly more frequently (an average of 13.6% of trials) than young preschoolers, $t_{166} = 4.80$, $p < .001$. Older preschoolers used *surprise* for the 'surprise' face (66.7%) on average significantly more frequently than the young preschoolers did (28.6%), $t_{166} = 5.51$, $p < .001$, but also used *surprise* marginally more frequently for the 'fear'

face (14.3%) than the young preschoolers did (6.5%), $t_{166} = 1.78$, $p = .08$ A similar pattern was observed for *fear* and *disgust*.

2.4. The categorization task: extending the Differentiation Model

We next turn to the categorization task in order to explore questions that could not be answered with the production task of free labeling.

2.4.1. Labeling Level and categorization

The first question is whether Labeling Level (as determined from the free labeling task) predicted performance on the categorization task. In this analysis, we used the subsample of 131 children who fit the Differentiation Model and had been assigned a Labeling Level in the free labeling task (i.e., this analysis omitted the 37 children who did not fit the pattern of Fig. 2; when these children were assigned the most plausible Labeling Level, the results were similar). Of course, Labeling Level and age were correlated, $r = .62$, $p < .001$, and thus, if one of them predicts performance on the categorization task, the other can be expected to do so. The question addressed here is whether age and Labeling Level independently predicted a significant amount of the variance in the number of "correct" responses on the categorization task. First, Labeling Level (seven levels: 0, 1, 2 (including Labeling Levels 2a and 2b), 3, 4 (including Labeling Levels 4a and 4b), 5, 6) and age (in months) were entered into a regression equation, with total correct (max = 32) on the four box trials as the dependent variable. With both age and Labeling Level in the equation, the variance accounted for was significant, $R^2 = .42$, $F(1, 128) = 46.69$, $p < .001$. When Labeling Level was entered first in a stepwise regression, the variance accounted for in this first step was significant, $R^2 = .35$, $F(1, 129) = 70.19$, $p < .001$, and age accounted for an additional 6.9% of the variance, which was also significant although not great, $\Delta R^2 = .07$, $F(1, 128) = 15.37$, $p < .001$. Conversely, when age was entered first, the variance accounted for in the first step was significant, $R^2 = .33$, $F(1, 129) = 63.10$, $p < .001$, and Labeling Level accounted for an additional 9.3% of the variance, which was also significant although not great, $\Delta R^2 = .09$, $F(1, 128) = 20.66$, $p < .001$.

2.4.2. Breadth of different categories

In the current and previous free labeling data, later-emerging categories, such as fear, appeared to begin narrower than happiness, anger, or sadness. Next we explored the breadth of children's emotion categories, as revealed in the categorization task by the proportion of faces that children included in each box. To keep this analysis parallel to the free labeling analyses, the 'excitement' and 'contentment' faces were excluded. Thus, only the six basic-level facial expressions were counted for each box. This analysis showed a different pattern than had been seen in the free-labeling data. In this comprehension task, the fear and anger boxes/categories were the broadest, followed by happiness and sadness. All comparisons were conducted using dependent samples *t*-tests. Children on average included as many faces in the later-emerging fear box/category (.46) as in the anger box/category (.41), $t_{167} = .53$, $p = .60$, and significantly more than they included in the happiness box (.34), $t_{167} = 6.25$, $p < .001$, or in the sadness box (.33), $t_{167} = 7.55$, $p < .001$. The same pattern held for the proportion of nontarget faces ('errors') included in each box: Children included as many nontarget faces in the fear box (.40) as in the anger box (.37), $t_{167} = 1.30$, $p = .20$, and significantly more than they included in the happiness box (.21), $t_{167} = 7.71$, $p < .001$ or in the sadness box (.22), $t_{167} = 8.55$, $p < .001$. But the original pattern observed in free labeling occurred in children's inclusions of the target faces: Children significantly more often included the target

face in the happiness box (.96), $t_{167} = 4.33$, $p < .001$, the anger box (.90), $t_{167} = 2.29$, $p = .02$, and the sadness box (.89), $t_{167} = 1.08$, $p = .03$, than in the fear box (.82).

2.4.3. Valence of children's 'errors' on the categorization task

We next examined the role of valence in the 'errors' seen in the categorization task. For each negative box, there were two mutually exclusive types of inclusion 'errors': (1) faces of the same valence as the target (e.g., for the sadness box, 'anger,' 'disgust,' and 'fear'), and (2) faces of opposite valence (e.g., for the sadness box, 'excitement,' 'happiness,' and 'contentment'). Because including the 'excitement' and 'contentment' faces in the happiness box cannot be considered 'errors' in the same way that including, for example, the 'anger,' 'disgust,' and 'fear' faces in the sadness box can be, we excluded the happiness box. Because there were only two children at Labeling Level 2b, those children and that Level were not included in these analyses. In a weighted means, repeated measures ANOVA (alpha = .05), Labeling Level (eight levels: Labeling Levels 0, 1, 2a, 3, 4a, 4b, 5, 6) was the between-subject factor; the three boxes (sadness, anger, fear) and two inclusion 'errors' (same valence, opposite valence) were within-subject factors. The dependent variable was the proportion of inclusion errors.

The results are shown in Table 1. The main effect for 'error' type was significant, $F(1, 121) = 101.29$, $p < .001$, with more same-valence than opposite-valence errors. There was also a main effect for Labeling Level, $F(7, 121) = 10.61$, $p < .001$ (Table 1), with fewer errors at higher Labeling Levels. Both of these main effects are qualified by the significant 'error'-type × Labeling Level interaction, $F(1, 121) = 2.11$, $p = .047$ (Table 1). Three observations can be made based on these results. First, inclusion of faces of the opposite valence decreased significantly and early. Fischer's LSD comparisons indicated that children at Labeling Level 0 included significantly more of the opposite-valence faces ($p = .03$) than did Labeling Level 1; Labeling Level 1 included significantly more ($p = .03$) than did Labeling Level 3. By Labeling Level 3, inclusions of the opposite valence faces had reached a floor level. Second, inclusions of faces of the same valence decreased significantly, but only gradually. LSD comparisons indicated that children at Labeling Levels 0 and 1 did not differ in the proportion of same valence faces they included in the boxes.

Table 1
Proportion of opposite- and same-valence errors on the categorization task

Labeling Level	Valence of inclusion error type		Mean errors
	Opposite valence	Same valence	
0	$.65_{a}$	$.81_{a}$	$.73_{m}$
1	$.31_{b}$	$.78_{a}$	$.54_{m}$
2a	$.11_{bd}$	$.48_{c}$	$.29_{no}$
3	$.04_{d}$	$.45_{c}$	$.25_{n}$
4a	$.10_{bd}$	$.35_{ce}$	$.23_{nop}$
4b	$.05_{d}$	$.30_{be}$	$.17_{nop}$
5	$.03_{d}$	$.25_{be}$	$.14_{op}$
6	$.00_{d}$	$.17_{be}$	$.09_{p}$
Mean	.09	.37	

Note. Maximum possible is 1.00. Proportion of same-valence errors is proportion of negative faces other than the target included in the box; proportion of opposite-valence errors is proportion of excitement, happiness, and contentment included. Least significant difference (LSD) comparisons of proportions (alpha = .05) were calculated on the mean proportion inclusions types. Means in the same row that do not share a subscript differ at $p < .001$. Means in the same column that do not share a subscript differ at $p \leq .05$.

Children at Labeling Levels 0 and 1 included significantly more ($p \leq .05$) of these faces than did children at Labeling Levels 2a and higher. Children at Labeling Levels 2a, 3, and 4a did not differ significantly; Labeling Level 3 included significantly more ($p < .03$) of these faces than did Labeling Levels 4b, 5, and 6. Labeling Levels 4a, 4b, 5 and 6 did not differ in the proportion of same-valence faces included in the boxes.

2.4.4. Narrowing of emotion categories with age and Labeling Level

We first examined the effects of age on children's emotion categories. In four parallel repeated measures ANOVAs (alpha = .05), one for each categorization trial, age (two levels: young, older) and gender (two levels) were between-subjects factors, and the eight facial expressions (surprise, excitement, happiness, contentment, sadness, disgust, anger, fear) were the within-subject factor. The dependent variable was whether, or not, each face was included in the box, scored 1 or 0, respectively.

In all four ANOVAs, the main effects for age and facial expression were significant, but these effects are qualified by the facial expression × age interaction, which was also significant in all four analyses: Happiness box, $F(4, 1148) = 7.83$, $p < .001$; Anger box, $F(4, 1148) = 8.74$, $p < .001$; Sadness box, $F(4, 1148) = 6.23$, $p < .001$; Fear box, $F(4, 1148) = 8.45$, $p < .001$. The results for all four boxes reveal a similar pattern: First, most of the children in both age groups included the target face (see Fig. 4). Second, older children were significantly less likely to

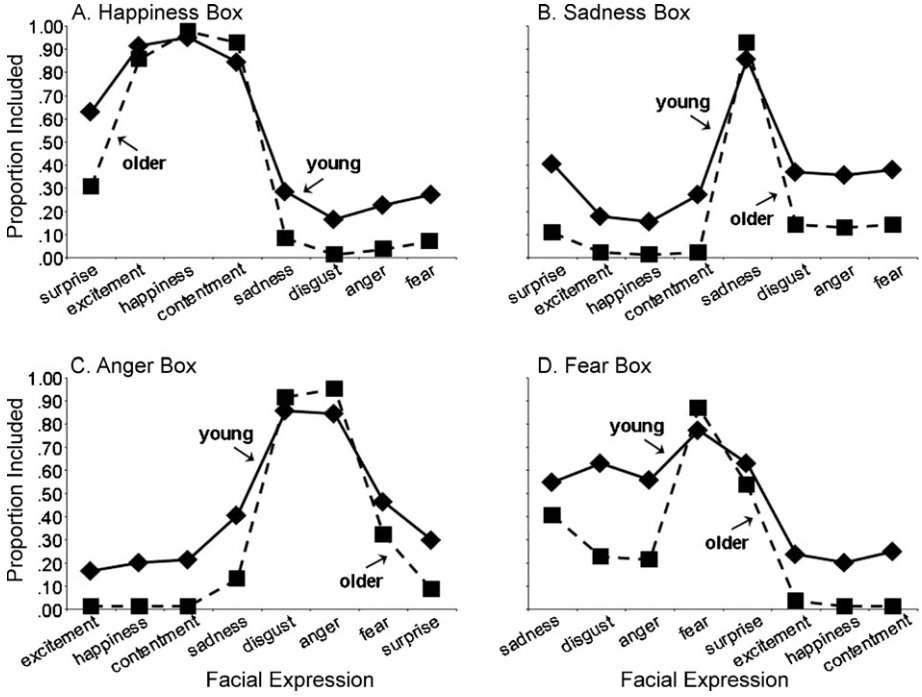

Fig. 4. The proportion of children in each age group who included each facial expression in (A) the Happiness Box, (B) the Sadness Box, (C) the Anger Box, and (D) the Fear Box. The facial expressions for each box are ordered in accordance with a structural model based on similarity of pleasantness and arousal, but the order is rotated for the Anger and Fear boxes.

include nontarget faces than younger children, thus showing a narrowing of emotion categories with age. In the happiness box (Fig. 4A), these differences were significant for surprise, sadness, disgust, anger, and fear faces (LSD comparisons, $p < .05$); in the sadness box (Fig. 4B), surprise, excitement, happiness, contentment, disgust, anger, fear faces ($p \leq .03$); in the anger box (Fig. 4C), surprise, excitement, happiness, contentment, sadness, fear ($p < .05$); and in the fear box (Fig. 4D), disgust, anger, and contentment ($p < .04$). Finally, in two of the boxes, both the younger and the older children included specific nontarget faces at the same rates as the target faces. In the happiness box, almost all the children included excitement and contentment faces, as well as the happiness faces; as many younger children included the happiness face as excitement face, but significantly fewer contentment face than happiness face ($p = .03$); as many older children included the happiness face as contentment face, but significantly fewer included the excitement face than the happiness face ($p = .01$). In the anger box, almost all the children included the disgust face as well as the anger face; there were no significant differences between the proportions of these two faces included by either age group. In addition, in the fear box, children in both age groups included the surprise face, although significantly fewer than included the fear face ($p < .01$). As many younger children included the surprise face as included the sadness, disgust, and anger faces; more tellingly, there was no significant difference between the proportion of older and younger children who included the surprise face, but significantly fewer ($p \leq .001$) older children included the sadness, disgust, or anger faces.

Next we tested whether the process by which emotion categories narrow (that is, exclude faces that do not belong) is more a function of age or of Labeling Level. With results from all four boxes/categories combined, the proportion of nontarget faces included was regressed on Labeling Level (seven levels: 0, 1, 2 (including Labeling Levels 2a and 2b), 3, 4 (including Labeling Levels 4a and 4b), 5, 6) and age (in months). With both age and Labeling Level in the equation, the variance accounted for was significant, $R^2 = .40$, $F(1, 128) = 43.24$, $p < .001$. When Labeling Level was entered first in a stepwise regression, the variance accounted for in this first step was significant, $R^2 = .35$, $F(1, 129) = 70.16$, $p < .001$, and age accounted for an additional 5.1% of the variance, $\Delta R^2 = .05$, $F(1, 128) = 10.93$, $p = .001$. Conversely, when age was entered first, the variance accounted for in the first step was significant, $R^2 = .30$, $F(1, 129) = 54.03$, $p < .001$, and Labeling Level accounted for an additional 10.8% of the variance, $\Delta R^2 = .11$, $F(1, 129) = 23.17$, $p < .001$.

We next examined the narrowing of children's emotion categories using Labeling Level, rather than age, as the measure of development. With Labeling Level we know when children have begun to use particular emotion labels on the free labeling task and thus can test various assumptions and predictions. We also begin to explore whether the rate of category narrowing changes around the time a particular label first occurs on the free labeling task. Specifically, does using a target label in free labeling (e.g., *surprise*, Labeling Level 4b) mark a sharp change in how children categorize the corresponding face (e.g., in the happiness box)? (Because there were only two children at Labeling Level 2b, those children and that Level were not included in these analyses).

Fig. 5A shows results for the happiness box. To simplify Fig. 5A, we grouped together results from the negative faces ('sadness,' 'anger,' 'disgust,' 'fear'); when they were examined separately, we found no indication of their being treated differently in the happiness box. Thus, Fig. 5A shows the proportion of participants who included the target faces ('excitement,' 'happiness,' 'contentment'), a comparison face ('surprise,') and opposite valence (negative) faces in the happiness box, plotted as a function of Labeling Level. Fig. 5B (Sadness Box), 5C (Anger Box), and 5D

(Fear Box) follow the same scheme. Each shows the proportion of participants who included the target face, opposite valence (positive) faces, a comparison face (except the Sadness Box). Each of these three boxes also shows the proportion of other negative faces (e.g., for the Sadness Box, the disgust, anger, and fear faces were grouped together).

Four observations can be made. First, for all four boxes, even children at the lowest Labeling Levels included the target face(s) in the box with high frequency. Second, at the earliest Labeling Levels, the children had begun to treat opposite-valence faces differently than the target face. So, differentiation based on valence had begun and was largely complete by Labeling Level 2. The one exception to this overall pattern may be the Happiness Box (Fig. 5A). Although at least half the children at the lower Labeling Levels excluded the negative faces, this result for negative faces could also be interpreted as the children here choosing randomly because they simply did not know if the negative faces were happiness.

Third, for the Sadness Box (Fig. 5B), Anger Box (Fig. 5C), and Fear Box (Fig. 5D), differentiation of the target face from the other negative faces was slower than for the positive faces. For the Sadness Box, results for the other negative faces were intermediate between these two extremes, at least until Labeling Level 4. For example, at Labeling Level 3, when children began using *sadness*, the proportion who included the other negative faces was significantly higher than the proportion who included the positive faces, dependent samples $t_{25} = 3.97$, $p < .001$. For the Anger Box, differentiation of anger from the other negative faces was slower. It began around Labeling Level 2, but proceeded more slowly than in the sadness box, and may not have been complete even at the highest Labeling Level for the preschoolers. For the Fear Box, differentiation

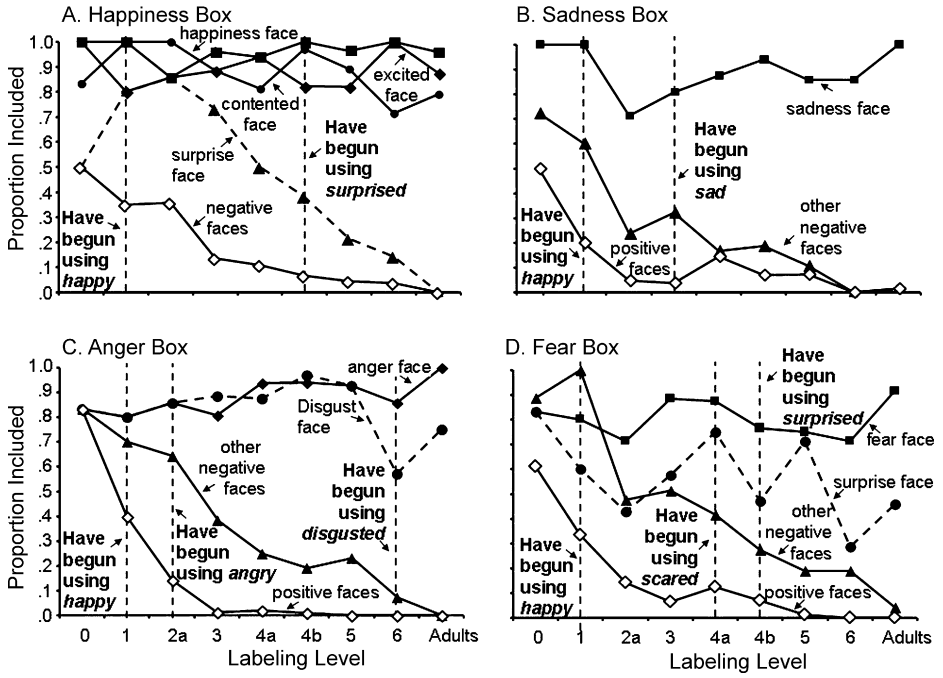

Fig. 5. The proportion of children at each Labeling Level who included each face in (A) the Happiness Box, (B) the Sadness Box, (C) the Anger Box, and (D) the Fear Box.

of the 'fear' face from the other negative faces was slow, slower even than for the anger box: This finding is a second indication that the fear category develops later than the other three categories. Differentiation of the fear face from the other negative faces begins at Labeling Level 2 and shows the most gradual decrease of all four boxes.

Finally, for the Happiness Box (Fig. 5A), Anger Box (Fig. 5C), and Fear Box (Fig. 5D), differentiation of the target face from a specific other face was analyzed. For the Happiness Box, we looked at inclusions of the surprise face, with a particular focus on Labeling Level 4b, when children begin using *surprise* in free labeling, to see if it would affect their categorization of this face as happiness. Only gradually over the full course of 3.5 years (24–65 months), between Labeling Levels 1 and 6, did they narrow the *happiness* category to exclude the 'surprise' facial expression. There was no sharp change in children's performance when they began using *surprise* in free labeling. It is more difficult to interpret the results for the child at Labeling Level 0, but perhaps these children were choosing randomly for the 'surprise' face.

For the Anger Box (Fig. 5C), we looked at inclusions of the disgust face, with a particular focus on Labeling Levels 2a and 6, when children were first observed to use *anger* and *disgust* in free labeling, to see if it would affect their categorization of this face as anger. Children did not distinguish between the 'anger' and 'disgust' faces until Labeling Level 6. At Labeling Levels 0 to 5, when children do not use *disgust* during free labeling, as many children included the 'disgust' face in the anger box as included the anger face. But at Labeling Level 6, when children were first observed to use *disgust* in free labeling, the proportion of children who included the 'disgust' face in the anger box decreased significantly from Labeling Level 5 ($t_{33} = 2.57$, $p = .01$) and was substantially lower (but not significantly, dependent samples t-test, $p = .17$) than the proportion of children who included the anger face at Labeling Level 6. Nevertheless, a majority of Labeling Level 6 children still included the 'disgust' face in the anger box. (Indeed, the adult comparison group's inclusion of the 'disgust' face suggests that Labeling Level 6 children's performance was nearly adult-like).

For the Fear Box (Fig. 5C), we looked at inclusions of the surprise face, with a particular focus on Labeling Levels 4a and 4b, when children were first observed to use *fear* and *surprise* in free labeling, to see if it would affect their categorization of this face as fear. For the 'surprise' face, which is adjacent to the 'fear' face on the circumplex model (Fig. 1), children's inclusions show an overall downward trend, but decreases in inclusions with Labeling Level are not as regular as for the other faces. Perhaps children's difficulty excluding the 'surprise' face from the fear category is more understandable when one considers that 47% of the adults included the 'surprise' face in the fear box. Thus, there was again no sharp change in children's categorization when they were first observed to use *fear* in free labeling.

3. Discussion

Both the free labeling and categorization data support our previous work (Widen & Russell, 2003). Moreover, for both tasks, results demonstrated the value of analyzing all responses, both 'correct' and 'incorrect.' This study also extended the Differentiation Model by testing whether it could predict children's performance on another emotion task. Indeed, it did: Children's Labeling Level within the Model predicted a significant amount of variance on the categorization task, beyond that accounted for by age. Thus, this study is the first to show that the Differentiation Model has predictive value for children's performance on another emotion task. Of course, the cross-sectional nature of the current study limits the strength of the conclusions we can draw

from the results. Stronger conclusions could be drawn from a longitudinal study, which we are currently conducting.

The Differentiation Model was originally based on children's first observed use of emotion labels on the free labeling task (Widen & Russell, 2003). The current data clarify that this first observed use is not a major milestone, even though it does provide useful information. Like many other acquisitions, the growth of emotion concepts (as measured here by children's use of emotion labels and their categorization of facial expressions) is gradual, yet gives rise to either/or milestones—somewhat arbitrary but highly visible events that index the underlying changes.

On the categorization task, children's later-emerging emotion category (fear) was as broad as their early-emerging ones (happiness, sadness, anger). This important finding corrected a mistaken interpretation from free labeling. In our original formulation of the Differentiation Model, based on children's free labeling responses to facial expressions, we proposed that children's early emotion categories (happiness, sadness, anger) were broader than their later-emerging categories (fear, surprise, disgust) (Widen & Russell, 2003). The categorization/comprehension task used here showed instead that fear (and presumably other later-emerging categories) is initially just as broad. Later-emerging categories may have appeared to be narrower in free labeling because label accessibility plays a role on this task. Emotion categories are fuzzy, even for adults (Russell & Bullock, 1986), making the role of accessibility more central for children. In free labeling, the earlier-emerging category labels are more accessible because they are more practiced. When a given face fits into more than one category, the more accessible label is more likely to be used. The labels for early emotion categories are used more frequently, especially by younger children (Wellman et al., 1995), and are thus more accessible and more likely to be used to label a variety of emotion stimuli, thus resulting in the apparent greater breadth of the early emotion categories. Once the production demands were removed in the comprehension task, it became clear that all four of the emotion categories we tested were equally broad.

The comprehension task showed that children's emotion categories narrow gradually. Narrowing began before children used a target label in free labeling and continued to be gradual after they used that label. This gradual narrowing is in accordance with other evidence that children's emotion categories develop gradually (Bormann-Kischkel, Hildebrand-Pascher, & Stegbauer, 1990; Denham & Couchoud, 1990; Russell & Widen, 2002b; Widen & Russell, 2003). Thus, the first observed use of an emotion label on free labeling is just one step along the way to understanding an emotion concept. This result supports the view that learning words is a part of learning concepts and that word learning and concept learning develop together (Gopnik & Meltzoff, 1997). One is not the prerequisite of the other.

One intriguing exception to this overall pattern was children's use of disgust in free labeling and subsequent categorization of the 'disgust' faces as not-angry. Within the set of emotion categories studied here, disgust is the last to emerge. In other research, disgust contrasts with other emotion categories. For example, in studies that contrast the power of an emotion label (such as *disgust*) with a facial expression (such as the prototypical "disgust facial expression") as a cue in eliciting a child's concept of that emotion, disgust provided the strongest contrast: for disgust, the label showed the greatest superiority over the facial expression (Camras & Allison, 1985; Russell & Widen, 2002a; Widen & Russell, 2004). One possibility worth exploring is that, for disgust, learning the label plays an especially important role in acquiring the concept. This hypothesis is in keeping with a renewed interest in the role of language in concept acquisition in general (Gelman, 2003; Gentner & Goldin-Meadow, 2003), and in the role of language in emotion understanding in particular (Lindquist et al., 2006).

In this study, we used the first appearance of an emotion label on the free labeling task as a marker of a child's having acquired the label as part of their productive vocabulary. This marker was possibly too conservative and therefore misleading. Indeed, the children in our study must have acquired emotion labels earlier than revealed by the free labeling task. Evidence from observational studies suggest that children begin using emotion labels even before their second birthday (Bretherton & Beeghly, 1982) and are already using a variety of emotion labels before their third birthday (Ridgeway et al., 1985; Wellman et al., 1995). Given that children use emotion labels earlier than revealed by the task of labeling facial expressions, it is also likely the children in the current study began using labels earlier than midway through the process of narrowing their emotion categories. For example, referring to Fig. 5C, suppose that all children, including those at Labeling Level 0, had acquired the label *anger* in some sense, perhaps having used the label spontaneously in conversation. What Fig. 5C would then show is even more extreme gradualness of the development of the anger category. In this scenario, label acquisition occurs earlier (indeed, perhaps it *is* the first step) in acquiring that emotion category, but the category is at that point at its broadest, as revealed by the categorization task. Once the label is acquired, several years of experience elapse as children's categories narrow and gradually become more adult-like. Future research might draw on parental reports of children's emotion vocabularies as an additional measure of the emotion labels children use.

One potential limitation of the current study is that both tasks used here involved children's understanding of facial expressions. Thus, a legitimate concern may be whether children's understanding of particular emotion concepts is changing or if their ability to identify facial expressions is changing. Although we used the best exemplars of facial expressions available, some properties of our results may depend on properties of the particular expressions used. Prior research has compared children's understanding of different aspects of emotion (e.g., faces versus labels; faces versus emotional events) (Camras & Allison, 1985; Russell, 1990; Russell & Widen, 2002a, 2002b; Widen & Russell, 2004). Researchers have asked preschoolers to choose from an array the face or label that matched stories describing emotional events, to describe the causes and consequences of emotions presented as either faces or verbal labels, and to categorize facial expressions when the target categories are specified by either the face or the label. In each case, results for the facial expression condition were similar to those for the label condition. Thus, we expect non-facial stimuli to yield results similar those obtained here. Nevertheless, it is important to test this prediction empirically.

The results of studies in which children were asked to describe the causes of different emotions also support the results of the categorization task more directly. These studies have found that preschoolers differentiate among the causes of negative emotions but have more difficulty differentiating among positive emotions. In describing the causes of different negative emotions, children rarely, for example, described anger causes for sadness (Dunn & Hughes, 1998; Russell, 1990). But in describing the causes of different positive emotions, although preschoolers could describe happiness, their descriptions of excitement and surprise were as likely to be judged as happiness stories as they were stories describing the target emotion (Russell, 1990). Thus, on both this story-telling task and the categorization task, preschoolers differentiated among the negative emotions but included all positive emotions in the happiness category.

The present findings have methodological implications. A typical method of investigating children's understanding of emotion is either to present them with emotion labels or to ask them to use the labels themselves. The current study shows that children's understanding of emotion

labels can be different from adults' even when they use the same labels. In addition, researchers often implicitly assume that if a child uses an emotion label in a study, whether the study is observational or experimental in nature, that child *has* the corresponding concept. The current study suggests that label acquisition is but one step in a years-long process toward understanding emotions in an adult-like manner.

The nature of children's early emotion concepts also has implications for childrearing and education. Young children may say they are *angry* but mean something quite different, something that fits within a much broader, umbrella category that they call *anger*. Conversely, young children may see a caregiver's nervous expression and interpret it as anger, causing them to feel that they have done something wrong. Understanding the nature of young children's emotion categories and their developmental paths should aid parents and teachers both to understand children and to help children understand emotions.

Acknowledgements

We thank everyone at the daycares that participated in this study: Without their help, this research could not be done. We thank Aimee Brooks, Diana Garrido, Maria-Paz Rodriguez, and Ryan Winton for their help with data collection and coding, and Mary Kayyal and Nicole Nelson for their input on the manuscript. We also thank Linda Camras for providing the photographs of facial expressions used in the free labeling task in this study. This study was funded by a grant from the National Science Foundation.

References

Bormann-Kischkel, C., Hildebrand-Pascher, S., & Stegbauer, G. (1990). The development of emotional concepts: A replication with a German sample. *International Journal of Behavioral Development*, *13*, 355–372.

Bornstein, M. H. (1985). On the development of color naming in young children: Data and theory. *Brain and Language*, *26*, 72–93.

Bretherton, I., & Beeghly, M. (1982). Talking about internal states: The acquisition of an explicit theory of mind. *Developmental Psychology*, *18*, 906–921.

Bretherton, I., Fritz, J., Zahn-Waxler, C., & Ridgeway, D. (1986). Learning to talk about emotions: A functionalist perspective. *Child Development*, *57*, 529–548.

Bretherton, I., McNew, S., & Beeghly-Smith, M. (1981). Early person knowledge as expressed in gestural and verbal communication: when do infants acquire a "Theory of Mind"? In M. E. Lamb & L. R. Sherrod (Eds.), *Infant social cognition* (pp. 333–373). Hillsdale, NJ: Erlbaum.

Bullock, M., & Russell, J. A. (1984). Preschool children's interpretation of facial expressions of emotion. *International Journal of Behavioral Development*, *7*, 193–214.

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.

Carey, S., & Bartlett, E. (1978). Acquiring a single new word. *Papers and Reports and Child Language Development*, *15*, 17–29.

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

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.

Dollaghan, C. (1985). Child meets word: "Fast mapping" in preschool children. *Journal of Speech & Hearing Research*, *28*, 449–454.

Dunn, J., Bretherton, I., & Munn, P. (1987). Conversations about feeling states between mothers and their young children. *Developmental Psychology*, *23*, 132–139.

Dunn, J., & Hughes, C. (1998). Young children's understanding of emotions within close relationships. *Cognition and Emotion*, *12*, 171–190.

Ekman, P., & Friesen, W. V. (1976). *Pictures of facial affect*. Palo Alto, CA: Consulting Psychologists Press.

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

Gelman, S. A. (2003). *The essential child: Origins of essentialism in everyday thought*. New York, NY: Oxford University Press.

Gentner, D., & Goldin-Meadow, S. (Eds.). (2003). *Language in mind*. Cambridge, MA: MIT Press.

Gopnik, A., & Meltzoff, A. N. (1997). *Words, thoughts, and theories*. Cambridge: MIT Press.

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

Heibeck, T. H., & Markman, E. M. (1987). Word learning in children: An examination of fast mapping. *Child Development*, *58*, 1021–1034.

Hollingshead, A. B. B., & Redlich, F. C. (1958). *Social class and mental illness: Community study*. Hoboken, NJ: John Wiley & Sons.

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.

Lindquist, K. A., Barrett, L. F., Bliss-Moreau, E., & Russell, J. A. (2006). Language and the perception of emotion. *Emotion*, *6*, 125–138.

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.

Matsumoto, D., & Ekman, P. (1988). Japanese and Caucasian facial expressions of emotion and neutral faces (JACFEE): 56 jpeg photos of facial expressions [CD]. http://www.paulekman.com/researchproducts.html.

Nelson, N. L., Widen, S. C., & Russell, J. A. (2006). *Children's understanding of emotions' causes and consequences: Labeling facial expression and stories*. Poster presented at the 18th Annual American Psychological Society Convention, New York, NY.

Rice, M. L. (1990). Preschoolers' QUIL: Quick incidental learning of words. In G. Conti-Ramsden & C. E. Snow (Eds.), *Children's language. Children's language series* (pp. 171–195). Hillsdale, NJ: Lawrence Erlbaum.

Ridgeway, D., Waters, E., & Kuczaj, S. A., III. (1985). Acquisition of emotion-descriptive language: Receptive and productive vocabulary norms for ages 18 months to 6 years. *Developmental Psychology*, *21*, 901–908.

Roberson, D., Davidoff, J., Davies, I., & Shapiro, L. R. (2004). The development of color categories in two languages: A longitudinal study. *Journal of Experimental Psychology: General*, *133*, 554–571.

Russell, J. A. (1980). A circumplex model of affect. *Journal of Personality and Social Psychology*, *39*, 795–804.

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

Russell, J. A., & Bullock, M. (1986). Fuzzy concepts and the perception of emotion in facial expressions. *Social Cognition*, *4*, 309–341.

Russell, J. A., & Paris, F. A. (1994). Do children acquire concepts for complex emotions abruptly? *International Journal of Behavioral Development*, *17*, 349–365.

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

Russell, J. A., & Widen, S. C. (2002b). A label superiority effect in children's categorization of facial expressions. *Social Development*, *11*, 30–52.

Sandhofer, C. M., & Smith, L. B. (1999). Learning color words involves learning a system of mappings. *Developmental Psychology*, *35*, 668–679.

Vicari, S., Reilly, J. S., Pasqualetti, P., Vizzotto, A., & Caltagirone, C. (2000). Recognition of facial expression of emotions in school-age children: The intersection of perceptual and semantic categories. *Acta Pædiatr*, *89*, 836–845.

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.

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. (2003). A closer look at preschoolers' freely produced labels for facial expressions. *Developmental Psychology*, *39*, 114–128.

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. (2004). The relative power of an emotion's facial expression, label, and behavioral consequence to evoke preschoolers' knowledge of its cause. *Cognitive Development*, *19*, 111–125.

Wolf, D. P., Rygh, J., & Altshuler, J. (1984). Agency and experience: Actions and states in play narratives. In I. Bretherton (Ed.), *Symbolic play: The development of social understanding* (pp. 195–217). New York: Academic Press, Inc.