Title
Sorting Out Emotions: How Labels Influence Emotion Categorization

Journal
Developmental Psychology, 58(9)

ISSN
0012-1649

Authors
Price, Gwendolyn F
Ogren, Marissa
Sandhofer, Catherine M

Publication Date
2022-09-01

DOI
10.1037/dev0001391

Sorting out emotions: How labels influence emotion categorization

Gwendolyn F. Price (ORCID: 0000-0002-2019-0732)[1]
Marissa Ogren (ORCID: 0000-0003-4441-5219)[2]
Catherine M. Sandhofer (ORCID: 0000-0002-0488-7185)[1]

[1]University of California, Los Angeles
[2]Rutgers University, Newark

Correspondence concerning this article should be addressed to Gwendolyn F. Price, University of California, Los Angeles, Box 951563
Contact: gprice@ucla.edu

Acknowledgments
This work was supported by NIH grant F31-HD100067 to MO. The authors would like to thank the parent and child participants and members of the UCLA Language and Cognitive Development Lab.

Author Note:
The design, sample size, and analyses for the studies in this article were all pre-registered through the Open Science Framework (https://osf.io/8zxtp/?view_only=2858ccad34964c7eb7fafc95c870f8c3). Materials for these studies are available from the NimStim database (stimuli: Tottenham et al., 2009) or by emailing the corresponding author (all other materials). All data and code are also available by emailing the corresponding author.

Abstract

The ability to categorize emotions has long-term implications for children's social and emotional development. Therefore, identifying factors that influence early emotion categorization is of great importance. Yet, whether and how language impacts emotion category development is still widely debated. The aim of the present study was to assess how labels influence young children's ability to group faces into emotion categories for both earliest-learned and later-learned emotion categories. Across two studies, 128 2- and 3-year-olds (77 female; Mean age=3.04 years; 35.9% White, 12.5% Multiple ethnicities or races, 6.3% Asian, 3.1% Black, and 42.2% not reported) were presented with three emotion categories (Study 1=happy, sad, angry; Study 2=surprised, disgusted, afraid). Children sorted 30 images of adults posing stereotypical expressions into one of the three categories. Children were randomly assigned to either hear the emotion labels prior to sorting (e.g., "*happy* faces go here") or were not given labels (e.g., "faces like *this* go here"). Results revealed a significant effect of condition for Study 2, such that labels led to improved emotion categorization for later-learned categories ($F(1,60)=8.15$, $p=.006$, $\eta_p^2=0.024$). However, there was no significant effect of condition for the earliest-learned emotion categories in Study 1 ($F(1,60)=0.94$, $p=.337$, $\eta_p^2=0.013$). Taken together, these results suggest that labels are important for emotion categorization, but the impact of labels may depend on children's familiarity with the emotion category.

Keywords: Emotion, Categorization, Early Childhood, Development, Language

Sorting out emotions: How labels influence emotion categorization

Learning emotion categories is a critical aspect of social-cognitive development (e.g., Izard et al., 2001; Voltmer & Von Salisch, 2017). Emotion categories broadly refer to groups of emotional cues (e.g., facial configuration, body posture, tone of voice) which indicate how someone may be feeling. Because emotional cues exist on a continuum (e.g., slight smirk to broad smile) and no cue maps to an emotion 100% of the time (e.g., people may or may not scowl when angry; Barrett et al., 2019), children must learn where to draw boundaries between emotion categories to make inferences about someone else's emotional state. Understanding emotions allows insight into others' goals and behaviors (Reschke, et al., 2017), inference of others' feelings and prediction of others' likely behaviors (e.g., Wu et al., s2017; Olson, et al., 1988). Thus, understanding how children learn emotion categories is important. Recent theoretical work suggests that emotion categories may be constructed with the help of emotion words (Barrett, 2017; Hoemann, et al., 2019; Lindquist, et al., 2015), as words have been shown to help children learn about other social categories (Rhodes, et al., 2018). However, the specific role of emotion words in learning emotion categories deserves greater attention. To address this, we examined whether the presence of emotion words would affect emotion categorization in a sorting task.

Emotion categories are complex. This is because emotions are both socially based and abstract. That is, emotions are dynamic (Hoemann, et al., 2017), and involve multiple components (e.g., facial movements, tone of voice, contextual information, body posture, knowledge of how this person tends to react) which are continually changing across time. Additionally, emotional information is highly variable (Barrett, et al., 2019), such that the same emotion appears in visually distinct ways (e.g., "fear" might present differently while running from a dog or freezing from stage fright). Moreover, the same expression can be indicative of different emotions depending on the context (e.g., scowling in anger versus concentration).

Integrating all of this information to form emotion categories represents a significant challenge for children (Hoemann, et al., 2019).

Despite the complexity, children begin learning about emotion categories from young ages. Starting as young as the first year of life, infants discriminate certain emotion cues such as facial configurations (Schwartz et al., 1985) and tones of voice (Walker-Andrews & Grolnick, 1983) and emotion category formation begins to emerge based on these cues (for a review see Ruba & Repacholi, 2019). At age 14- to 18-months, when infants typically are not producing much general language or emotion language, research has shown that novel labels help infants to form superordinate emotion categories (Ruba et al., 2020b). This may be because even before infants are verbal and can produce emotion words themselves, receptive language may be helpful for identifying similarities across emotion categories (Shablack et al., 2020). By 18 months of age, children begin to produce some emotion words themselves such as “happy” (Ridgeway, et al., 1985), although other emotion words such as “disgusted” do not emerge in children’s speech until closer to 4.5 years of age. Children continue to learn new emotion words over the next several years, with emotion vocabulary rapidly increasing and becoming more adult-like from age 4 to 11 (Baron-Cohen et al., 2010; Grosse et al., 2021). Additionally, children’s ability to talk about emotions abstractly develops even more gradually, plateauing around age 18 (Nook et al., 2020). Thus, the ability to use emotion labels, and talk about emotion categories is complex and continues to develop across infancy, childhood, and adolescence.

Further research has shown that children tend to follow a common developmental progression in their understanding of emotion categories, with earliest learning of the categories “happy”, “angry”, and “sad” compared to later-learned categories such as “scared”, “surprised”, and “disgusted” (Widen & Russell, 2003). Children’s understanding of emotion categories develops rapidly through at least 5 years of age, with emotion categories beginning broad and gradually narrowing with development (Widen & Russell, 2008a). By 2 years of age, children

begin to map emotion labels to specific facial expressions in standardized tasks (e.g., the Affective Knowledge Test; Denham, 1986). Although emotions are inherently complex, children begin to learn them early on, so it is important to understand what factors may facilitate this learning. In particular, we examine the role of emotion words in children's emotion categorization. Because language may have a differential impact depending on when in development these emotion categories are typically learned, we examine this potential impact of language for earliest- vs later-learned emotion categories.

There are a number of reasons to expect that emotion words may help children to categorize emotions. First, research has shown that labels help children form a wide variety of categories (e.g., Fulkerson & Haaf, 2003; Gentner, et al., 2011; Johanson & Papafragou, 2016; Russell & Widen, 2002). Further, labels aid in categorization even when category members do not share much perceptual similarity (Gentner & Namy, 1999; Johanson & Papafragou, 2016). One explanation for why labels aid in categorization is that labels link category instances together, which facilitates abstraction of the category structure (Erickson, et al., 2014; Gliozzi, et al., 2009; Plunkett, et al., 2008). Labels may do this in two ways: 1) Linking category members together by directing children's focus to the commonalities between examples (Althaus & Plunkett, 2016), and 2) Making distinctions *between* categories more concrete (Lupyan, et al., 2007). The benefit of labels for category learning has been particularly demonstrated for relational categories, which are complex and abstract (Loewenstein & Gentner, 2005). Because emotions are also abstract and complex, labels may provide a similar advantage for learning emotion categories. Some research done with novel emotion categories has shown that children will use a process of elimination to identify new emotion categories when given an unfamiliar label (Nelson & Russell, 2016a, 2016b; Nelson et al., 2018). This notion is also supported by evidence showing that labels help adults to perceive categories of chimpanzee facial expressions (Fugate et al., 2010). Additionally, research with 6- to 25-year-olds has demonstrated that the relation between age and a more mature conceptualization of emotions

was mediated only by verbal ability (Nook et al., 2017). Thus, labels may also aid in emotion category learning.

In line with this perspective, some theoretical work has suggested that labels help children to develop emotion categories (e.g., Lindquist & Gendron, 2013). The Theory of Constructed Emotion (Barrett, 2017), for example, states that emotions are built from learning and experience, and that labels help to construct the categories of emotion in much the same way they do for other abstract concepts. Even in adults, hearing emotion labels while viewing emotional faces can impact the perception of that emotion, such that hearing “furious” versus “grumpy” changes the level of arousal from the image (Barker et al., 2020). It seems likely that emotion labels may affect children’s development of emotional categories, because children who learn emotion labels early on can categorize based on a label alone. For example, one study found that children were better able to sort based on emotion labels than emotional facial expressions (Russell & Widen, 2002).

However, other theoretical perspectives do not suggest a role for language in emotion understanding development. For example, Basic Emotions Theory suggests a strong evolutionary and biological basis for emotion understanding (Ekman, 1992). Basic Emotions Theory holds that language is a communication device for labeling emotions but is not crucial for developing an understanding of these emotions (Ekman & Cordaro, 2011). Yet, much remains unknown about which of these theoretical perspectives best explains the role of language in emotion categorization development. In line with the Theory of Constructed Emotion account, it has previously been argued that “...emotion words hold the key to understanding how children learn emotion concepts in the absence of biological fingerprints and in the presence of tremendous variation” (Barrett, 2017, p. 102). Barrett (2017) further states “So far, my hypothesis about emotion words is only reasoned speculation because the science of emotion is missing a systematic exploration of this question” (p. 102). Thus, additional research examining emotion sorting at an age in which children are rapidly developing the ability

to categorize emotions is crucial, as it may provide insight into the role of language in early emotional development, as well as clarify which theoretical perspective may most accurately explain early emotional development.

The Present Study

To examine whether emotion labels help children categorize emotional expressions we presented 2- and 3-year-old children with a sorting task. Children identified which category (depicted by an emoji) a real person's face belonged to. Emojis were selected because they present schematic versions of faces, allowing for a more abstract matching task and preventing children from matching content based on specific features of real faces. Previous research has used emojis in categorization paradigms with adults (Ruba, et al., 2018), and developmental studies have found that schematic, cartoon stimuli are easier for infants to associate with novel labels than real faces (Ruba, et al., 2020a). Thus, tasking children with sorting emotional faces into emoji categories with or without labels (but all other perceptual information the same) allows us to examine how the presence of labels may or may not influence the way that children categorize emotional content. Study 1 asked children to sort happy, sad, and angry faces. Study 2 asked children to sort surprised, disgusted, and afraid faces. These six emotion categories were selected because they are the first six emotion categories typically learned by children; *happy*, *sad*, and *angry* are typically learned earliest in development, followed by *disgusted*, *surprised*, and *afraid* (Widen & Russell, 2003). During the task, children were asked to sort faces into categories while hearing either emotion labels or no labels. Evidence in support of Basic Emotions Theory would be found if labels did not influence sorting performance, as this theory suggests that language should not influence sorting and understanding of these basic, evolutionarily important emotion categories. In contrast, if hearing emotion labels benefits emotion categorization, this would align more with a Theory of Constructed Emotion perspective (Barrett, 2017).

Study 1

To determine whether emotion labels affect categorizing the earliest-learned emotions, we examined whether presenting labels influenced sorting happy, sad, and angry faces. Parents report that children typically understand these words around 18-29 months (Ridgeway, et al., 1985), indicating that these words were likely familiar to our participants. We hypothesized that the 2-year-olds would benefit more from hearing emotion labels than the 3-year-olds because the 2-year-olds are still learning these categories whereas the 3-year-olds may already know the emotion words and categories and therefore may perform well regardless of the presence of emotion labels.

Method

Participants

Sixty-four 2- and 3-year-old children (32 2-year-olds; 34 female) participated in this study (11 White, 3 Multiple ethnicities or races, 1 Asian, 1 Black, 48 did not report). Racial information was collected only for online participants, and no SES data was collected. This sample size was selected in order to be equal to or slightly higher than sample sizes from previous studies using similar methodologies (Nelson & Russell, 2016a; Nelson et al., 2018), and was included in our pre-registration, which is described below. We further conducted a sensitivity analysis (Perugini et al., 2018) in G*Power 3.1 (Faul et al., 2009), and this confirmed that our sample size was large enough to detect a medium-large interaction effect (f=0.36) assuming an alpha level of .05 and power of 0.8. As some data were collected in local preschools, parents were not asked to provide SES information. An additional 16 children (all 2-year-olds) were excluded from the final dataset due failure to pass the familiarization trials (*N*=14) (as described below) or to an inability to sit through all 30 trials (*N*=2). Children were recruited through preschools, a birth records database, and an online recruitment site (childrenhelpingscience.com). The average participant age for the 2-year-olds was 2.60 years (*SD*=0.28) and for the 3-year-olds was 3.55 years (*SD*=0.28). Children received a book or $5 Amazon gift card for their participation for studies completed in person or online, respectively. This study was approved by the university ethics

committee (IRB approval #10-001578), and all parents provided written or verbal informed consent.

Data collection began in-person in a university lab space and at local preschools within [Removed for blinded review] county. However, with the onset of the COVID-19 pandemic, data collection converted to an online format for the remaining participants all located in the United States. Across both formats, data were collected between August 2019 and October 2020. 47 children participated in the study in-person, and the remaining 17 children participated online. Comparisons between in person and online participants are examined and discussed in the results section.

Stimuli

Thirty images from the NimStim facial expression database (10 happy, 10 angry, and 10 sad) were used (Tottenham et al., 2009). These pictures were designed to provide stereotypical examples of emotion categories as agreed upon by adults. We selected a subsample of the pictures in the NimStim database (5 men and 5 women images for each emotion) and which all included open-mouthed expressions to control for the salience of this feature. The same individual from the NimStim database did not repeat, so the 30 faces represented 30 unique individuals. The selected photos had a high average identification rating in the original validation study (happy = 99.1%, angry = 97.9%, sad = 83.1%), all substantially above what would be expected by chance. All children saw the same 30 validated facial expressions.

In-Person Procedure

Familiarization. Children first participated in a familiarization game to make sure they understood the rules of the game. To begin, three buckets were placed in a row in front of the child. Each bucket had an image of an emoji animal on it (i.e., dog, cat, and horse). The children were told they were going to play a sorting game. The experimenter pointed to each of the emoji animals in turn and said “(Horse, Cat, Dog) pictures go in this bucket.” Then the participant was handed a realistic image of one of the animals and asked, “So where does this one go?”

Children were given two chances for each and had to correctly sort all three images to move on to the next part of the study. This ensured that children knew how to sort for the game, and also recognized that the emoji images were representations of real-life images.

Emotion Sorting. Each bucket had a picture of a 1.25 inch x 1.25 inch emoji depicting an emotional expression (i.e., happiness, sadness, and anger). The emojis were similar on all dimensions aside from emotional expression and could not be differentiated by race (i.e., they were yellow), gender (i.e., emojis were ambiguous), eye color (i.e., black), or size (i.e., equivalent). In this way, the only feature that differed among the emoji's and therefore could be relevant for sorting the pictures of real human faces was their emotional expression. Children were randomly assigned to one of two conditions (Label or No Label). In the Label condition, the experimenter pointed to each of the emoji faces in turn and said "<Emotion Word> faces go in this bucket." The emotion words used were "happy", "sad", and "angry". In the No Label condition, the experimenter instead said, "These kinds of faces go in this bucket." Then the participant was handed one of the 3 inch by 3 inch NimStim images and asked, "So where does this one go?" Children were not given any feedback except a neutral phrasing, e.g., "thanks!" After every 3 images, the experimenter repeated the sorting rules in both conditions. Children sorted 30 faces in total.

Post Test. After sorting all 30 images, children were asked to verbally identify the emotions in the 3 emojis. The experimenter pointed to each emoji and asked, "How does he/she feel?" making sure to match the gender to the participant. Children were given no feedback except a neutral phrasing as before, e.g., "ok!"

Online Procedure

Sessions were conducted over zoom. Instead of three buckets, the emojis were displayed in three areas of the screen (i.e., top-center, bottom-left, and bottom-right). The relative size of emojis to realistic images of faces was preserved across the in-person and

online testing formats (although the actual size each child saw in the online procedure depended on their screen size).

Familiarization. Children first were asked to point to each of the three animal emojis, and were prompted with a phrase “Can you point to the (cat, dog, horse)?” Then children were told “When you see a picture of a (cat, dog, horse) in the middle, I want you to point to the (cat, dog, horse) emoji on the outside. Ok?” Then one at a time the realistic images of the cat, dog, and horse were presented in the middle of the screen and the experimenter said: “Can you point to where this one goes?” Children were provided two chances for each animal and had to point to all three correct emojis before moving on to the emotion sorting phase of the study.

Emotion Sorting. The three animal emojis were then replaced by three face emojis that depicted an emotional expression: happy, sad, and angry. In the Label condition, the experimenter used the computer mouse to point to each of the emoji faces in turn and said “<Emotion Word> faces go here.” The emotion words used were: “happy”, “sad”, and “angry”. In the No Label condition, the experimenter instead said, “These kinds of faces go here,” for each emoji. In both conditions a NimStim image appeared in the middle of the screen and the participant was asked “Can you point to where this one goes?” Children were not given any feedback except a neutral phrasing, e.g., “thanks!” After every 3 images, the experimenter repeated the sorting rules.

Post Test. After sorting all 30 images, children were asked to identify the emotions in the 3 emojis. One of the emojis was displayed in the center of the screen at a time and the experimenter asked, “How does he/she feel?” making sure to match the gender to the participant. Children were given no feedback except a neutral phrasing as before, e.g., “ok!”

Transparency and Openness

The design, sample size, and analyses for the studies in this article were all pre-registered through the Open Science Framework (https://osf.io/8zxtp/?view_only=2858ccad34964c7eb7fafc95c870f8c3). Materials for these

studies are available from the NimStim database (stimuli: Tottenham et al., 2009) or by emailing the corresponding author (all other materials). All data and code are also available by emailing the corresponding author.

Results

Testing Format (In-Person vs Online)

We began by examining whether children's performance changed across the two testing formats (in-person versus online), as the data collection method changed partway through the study due to COVID-19 disruptions. To assess this, we specifically analyzed the 2-year-olds responses, as approximately half of this age group participated in person (N=15) and half online (N=17). Further, within each data collection method, 8 of the participants were in the label condition, and the remainder were within the no label condition (in person N=7; online N=9). There were no significant differences in emotion sorting between the in person and online testing groups ($t(30)=0.45$, $p=.659$, $d=0.16$). Based on these results we concluded that sorting did not significantly differ based on the testing format. Thus, subsequent analyses were conducted collapsing across testing format.

Descriptive Statistics

We first asked whether children sorted at levels above chance. There were 30 trials and each trial presented children with 3 categories (chance = 33.33%). Thus, we expected that by chance children would make 10 correct responses. Results revealed that on average (across all ages and conditions) children made 21.70 correct responses (SD=7.14) which was significantly above chance ($t(63)=13.12$, $p<.001$, $d=1.64$), indicating that overall children performed well in this study.

Language and Emotion Sorting

We next proceeded to our primary question of interest - did the presence of labels affect sorting? To answer this question, we conducted a 2 (age) x 2 (condition) between subjects ANOVA. Figure 1a shows children's mean number of correct choices in the emotion sorting

task. There was a significant main effect of age such that the 3-year-olds (*M*=25.13, *SD*=4.52) had a significantly higher number of correct sorting responses than 2-year-olds (*M*=18.28, *SD*=7.67), ($F(1,60)=18.74$, $p<.001$, $\eta_p^2=0.093$). However, there was no significant main effect of condition such that sorting did not differ between the label (*M*=22.47, *SD*=6.73) and no label conditions (*M*=20.94, *SD*=7.54), ($F(1,60)=0.94$, $p=.337$, $\eta_p^2=0.013$), and there was no significant interaction between age and label condition ($F(1,60)=0.54$, $p=.535$, $\eta_p^2=0.002$). Children's errors can help inform how they are constructing their categories. As such, confusion matrices depicting children's incorrect choices by age and condition can be found in the supplementary materials. Although we do not conduct analyses on these confusion matrices as they were reported post-hoc, it may be interesting to note potential differences in children's pattern of responding across label and no-label conditions, particularly for the 2-year-olds.

Bayesian Analysis

To complement our analyses above, we performed follow-up Bayesian analyses investigating the effect of child age, condition, and their interaction on child sorting using JASP (JASP Team, 2020). Bayesian analyses can be informative in conjunction with standard null hypothesis testing, as these analyses allow for a comparison of the relative evidence for the null and alternative hypothesis (Lakens et al., 2020), and the strength of evidence in either direction. Results revealed extreme evidence for the effect of age on sorting (BF10=297.06). Further, results revealed anecdotal evidence for the null hypothesis for the effect of condition (BF10= 0.35) and for the age by condition interaction (BF10= 0.48), suggesting that the null results reported above should be interpreted with caution (Jeffreys, 1961).

Post Test Trials

Lastly, we examined children's performance on the post-test trials. Figure 1b shows children's mean number of correct responses in this free labeling task. Children's verbal responses for how each emoji was feeling were coded as correct or incorrect, and the number

of correct responses across the three emotions was summed. Synonyms were coded as correct (e.g., both “angry” and “mad” were correct responses for the angry emoji), but more general affective information was coded as incorrect (e.g., “bad” or “not good”). We compared post-test performance across the two conditions to determine whether the labels presented during the task impacted children’s ability to provide the appropriate labels at post-test. Two children refused to provide verbal responses at post-test (one child in each condition), and they were excluded from post-test analyses. Results from the remaining 62 participants revealed that there was no significant difference in post-test performance ($t(60)=1.10$, $p=.277$), with comparable performance in the Neutral ($M=2.55$, $SD=0.62$) and Label ($M=2.71$, $SD=0.53$) conditions. Thus, children in general were relatively successful at labeling the happy, sad, and angry emojis at post-test, regardless of whether or not they heard these labels throughout the task.

Study 2

Study 2 asked whether labels would affect emotion categorization if children were tasked with sorting later-learned emotion categories which children are less likely to produce labels for (Widen & Russell, 2003). Thus, in Study 2, we examined whether emotion labels influence young children’s ability to categorize the emotion categories of disgust, surprise, and fear. Parents report that their children typically understand these words between 24-59 months (Ridgeway, et al., 1985), indicating that they may be less familiar to our participants. Further, prior research has shown that these “later-emerging emotion categories” narrow with child age (Widen & Russell, 2008b), as children begin with broad representations of these categories that they gradually narrow down to adult-like categories, and this category narrowing typically occurs later than for anger, happiness, and sadness (Widen, 2013; Widen & Russell, 2008b). We hypothesized that for these later-emerging emotion categories, labels would help 3-year-olds more than 2-year-olds. This may be because 2-year-olds are too unfamiliar with these specific emotion categories, whereas 3-year-olds may have just enough familiarity with these later-emerging categories to gain an advantage from the label support.

Participants

Sixty-four 2- to 3-year-old children (32 2-year-olds, 43 female) participated in this study (35 White, 13 Multiple ethnicities or races, 7 Asian, 3 Black, 6 did not report). None of the children had participated in Study 1, and we deliberately targeted the same sample size as Study 1. An additional 22 children (19 2-year-olds) were excluded due to an inability to pass the familiarization trials (N=14) or sit through all 30 test trials (N=8). The average participant age for 2-year-olds was 2.56 years (SD=0.29) and for 3-year-olds was 3.46 years (SD=0.31), and exactly half of the participants were 2-year-olds and half were 3-year-olds. Recruitment, compensation, IRB approval, and consent were the same as Study 1. Data collection began in-person at local preschools but was then converted to an online format due to the onset of the COVID-19 pandemic, with participants located across the United States. Across both formats, data were collected between March 2020 and January 2021. Two of the children participated in the study in-person, and the remaining 62 children participated online.

Stimuli

Thirty new images from the NimStim facial expression database (10 afraid, 10 disgusted, and 10 surprised) were used in Study 2 (Tottenham et al., 2009). The faces were selected to equally represent both genders. As with Study 1, the faces all included open-mouthed expressions, and the 30 stimuli were from 30 unique individuals. The selected photos had a high average identification rating in the original validation study (afraid = 85.6%, disgusted = 97.0%, surprised = 92.5%).

Procedure

Sixteen 2-year-olds and 16 3-year-olds were randomly assigned to the Label and the No Label conditions. The procedure for both the in-person and online formats were identical to Study 1 with three changes. First, the new NimStim faces depicted afraid, disgusted, and surprised categories. Second, three new emojis were used to represent the new emotion

categories. Finally, in the Labels condition, the experimenter now used the emotion terms "afraid", "disgusted", and "surprised".

Results

Descriptive Statistics

Results revealed that on average (across all ages and conditions) children made 16.09 (out of 30 possible) correct sorting choices (*SD*=5.00), which was significantly above chance (*t*(63)=9.76, *p*<.001, *d*=1.21). However, as expected, performance was worse than in Study 1 (*t*(126)=5.15, *p*<.001, *d*=0.91) indicating that Study 2 was more difficult for children.

Language and Emotion Sorting

A 2 (age) x 2 (condition) ANOVA was conducted to determine whether there were differences in task performance between the groups. Figure 2a depicts children's mean number of correct sorting choices. There was a significant main effect of condition such that children in the label condition (*M*=17.78, *SD*=4.74) made more correct sorting choices than children in the no label condition (*M*=14.41, *SD*=4.74), *F*(1,60)=8.15, *p*=.006, η_p^2=0.024). However, there was no significant main effect of age such that sorting did not differ between the 2-year-olds (*M*=15.38, *SD*=3.87) and 3-year-olds (*M*=16.81, *SD*=5.89), (*F*(1,60) =1.48, *p*=.229, η_p^2=0.02). Additionally, there was no significant interaction between age and label condition (*F*(1,60) =0.72, *p*=.401, η_p^2=0.012). As with Study 1, confusion matrices are presented in the supplementary materials. Here, it may be interesting to note potential differences in children's pattern of responding across label and no-label conditions, particularly for the 3-year-olds.

Bayesian Analysis

As with Study 1, we performed follow-up Bayesian analyses investigating the effect of child age, condition, and their interaction on child sorting using JASP (JASP Team, 2020). Results for Study 2 revealed moderate evidence for the effect of condition on sorting (BF10=5.51). Further, results revealed anecdotal evidence for the null hypothesis for the effect

of age (BF10= 0.44) and for the age by condition interaction (BF10= 0.54), suggesting that the null results above should again be interpreted with caution.

Post Test Trials

As with Study 1, we examined the number of correct verbal responses on the post-test trials based on children's study condition. Figure 2b shows children's mean number of correct responses in this free labeling task. Synonyms such as "scared" for afraid and "yucky" for disgust were coded as correct. Seven children refused to provide verbal responses at post-test (two in the label and five in the neutral condition). Results from the remaining 57 children revealed that there was a significant difference in post-test performance ($t(55) = 4.37$, $p < .001$), with better performance in the Label ($M = 1.40$, $SD = 1.10$) than No Label ($M = 0.37$, $SD = 0.55$) condition. Children in general had a harder time labeling disgusted, surprised, and afraid faces at post-test relative to the emotions in Study 1, but children in the Label condition were aided by the presence of these labels throughout the task. However, it is interesting to note that even the children in the Label condition, who had been presented with these exact labels along with the emojis 10 times each throughout the task immediately prior, still answered fewer than half of the post-test questions correctly on average. Despite children in Study 2 not readily providing accurate labels at post-test, the presence of labels throughout the task still significantly improved children's sorting performance relative to the no label condition.

Language and Emotion Sorting: Across Studies

To compare across studies 1 and 2, a 2 (age) x 2 (condition) x 2 (study) ANOVA was conducted. There was a significant main effect of age such that 3-year-old children (M=20.97, SD=6.69) made more correct sorting choices than 2-year-old children (M=16.83, SD=6.20) regardless of study or condition ($F(1,120) = 17.60$, $p < .001$, $\eta_p^2 = 0.128$). There was also a significant main effect of condition such that children in the label condition (M=20.13, SD=6.24) made more correct sorting choices overall than children in the neutral condition (M=17.67,

SD=7.06) regardless of study or age (F(1,120) =6.18, p=.014, η_p^2=0.049). There was a third main effect of study such that children in study 1 (M=21.70, SD=7.14) made more correct sorting choices than children in study 2 (M=16.09, SD=5.00) regardless of age or condition (F(1,120) =32.30, p<.001, η_p^2=0.212).

In addition, there was one significant interaction between age and study (F(1,120) =7.50, p=.007, η_p^2=0.059). This interaction shows that 3-year-old children made more correct sorting choices in study 1 (M=25.13, SD=4.52) than in study 2 (M=16.81, SD=5.89) regardless of condition (t(62)=6.33, p<.001). However, 2-year-olds were not significantly different in their sorting choices between study 1 (M=18.28, SD=7.67) and study 2 (M=15.38, SD=3.87) (t(62) =1.91, p=.060). There was no significant interaction between condition and age (F(1,120) =.006, p=.937, η_p^2<.001), between condition and study (F(1,120) =.872, p=.352, η_p^2=.007), nor was there a three way interaction between age, condition, and study (F(1,120)=1.19, p=.277, η_p^2=.010). These results suggest that while there are significant differences between studies overall, the pattern of the results was not necessarily significantly different between Study 1 and 2, as would be indicated by a three-way interaction.

General Discussion

Across two studies, we examined whether emotion labels affected 2- and 3-year-old children's emotion sorting choices. In Study 1 we found no effect of labels for the emotions happy, sad, and angry, but did find differences between 2- and 3-year-olds' performance. In Study 2, we found that hearing the emotion label affected sorting performance for the emotions disgusted, surprised, and afraid, and this was mirrored in the free labeling post-tests as well. This suggests that labels may play an important role in sorting emotion categories. Further, our Bayesian analyses confirmed moderate-extreme evidence for our significant results, suggesting the strength of the effects of age in Study 1 and condition in Study 2. However, our null results should be interpreted with caution given the results of the reported Bayesian analyses, which

leave open the possibility for significant effects to be revealed among larger samples.

The present results have implications for theories of emotional development. Although some theories have emphasized the importance of labels for children's developing emotion categorization (e.g., Barrett, 2017; Lindquist & Gendron, 2013), other theoretical perspectives dispute this claim. Such theories posit that the effect of labels does not extend to emotion categories because emotion expression and perception are heavily based in our evolutionary and biological history (Ekman, 1999). That is, emotions are evolutionarily based and therefore these labels are simply used to describe a state that we already know (Ekman, 1992). According to this perspective "Language and emotion are independent from each other" (Ekman & Cordaro, 2011) and therefore although emotion labels are used to communicate about emotion states, they may not be necessary for forming emotion categories. However, the results from the present study indicate that labels *do* influence how children categorize emotional information, and therefore support the Theory of Constructed Emotion (Barrett, 2017). Those who view emotion and language as independent might argue that labels only influenced emotion categorization because children were already familiar with these particular labels and emotion categories. However, the results of Study 2 demonstrated an advantage for emotion labels specifically when the emotion categories were less familiar. Moreover, children's overall performance was significantly lower in Study 2 than in Study 1, indicating that the labels were particularly helpful for categorizing the more challenging, later-learned emotions. Thus, the present results suggest that emotion labels aid children in sorting emotion categories and support theoretical perspectives which emphasize the role of language in emotional development.

Why might labels have aided emotion sorting in Study 2 (i.e., disgusted, surprised, and afraid), but not significantly in Study 1 (i.e., happy, sad, and angry)? Previous research indicates that children tend to learn emotion categories and emotion words, in a particular order (Ridgeway, et al., 1985; Widen & Russell, 2003), such that children typically produce the words

happy, *sad*, and *angry* before *disgust*, *surprise*, or *fear.* Both the 2- and 3-year-olds likely already had familiarity with the emotion categories happy, sad, and angry. However, the emotion categories disgust, surprise, and fear, were likely less familiar, as they tend to be later learned. One possibility is that any benefit of labels for emotion categorization may be more pronounced for emotion categories that children are still in the process of learning compared to categories they have already learned. This is because we used real emotion labels, and children may draw from both their prior experience with and understanding of this label along with the information in the task itself when determining how to categorize the faces. If this is the case, we may expect that labels would help children to sort happy, sad, and angry at ages younger than 2, when these categories are less familiar. Future research may wish to examine this possibility. In addition, future research may wish to examine if and how these results generalize to participants outside the US and in particular to languages other than English as languages present variation in emotional categories and labels. Thus, cross-linguistic differences may provide a way to examine how the ages at which children learn certain emotion categories relates to the influence of emotion labels.

Learning about emotions involves integrating many levels of verbal, perceptual, and conceptual meaning. Before children have mastered all the complexity involved in understanding emotion, labels may influence emotion understanding. That is, children may develop some mastery of emotion words, before they fully understand the deeper meanings of those words. For example, although children may learn an emotion word early in development (e.g., “anger”), children do not develop an adult-like understanding of the emotion (e.g., differentiating between images of “disgust” and “anger” and understanding the situations that may fit with those emotions) until much later (Widen, 2013). However, these partial understandings may contribute to children’s learning. Emotion labels may organize attention in the same way as other category labels (Hoemann et al., 2019; Shablack et al., 2020). One possibility is that the labels helped children to link the emojis to their prior knowledge about

emotions – drawing attention to the features of the emojis that characterize a particular emotion. There is evidence for this from other domains which demonstrate that labels draw attention to shared commonalties (Gentner & Markman, 1995; Waxman & Markow, 1995) aid in category discrimination (Lupyan, et al., 2007) and increase the perceived similarity between items that share a label (Sloutsky & Fisher, 2004). Thus, hearing and learning emotion labels is likely one step among many in a developmental cascade toward more fully understanding category concepts (e.g., Oakes & Rakison, 2019).

It is important to note that the present study included images of exaggerated, posed, stereotypical facial configurations, and thus may have differed from the rich contexts in which children perceive emotions in their everyday lives. We deliberately selected these images because a Basic Emotions perspective (Ekman, 1992) would hold that these emotion categories should not rely on language for emotion perception or sorting. However, emerging evidence suggests that emotion categories are highly variable, and that these stereotypical images are likely not representative of how emotions appear in day-to-day life (Barrett et al., 2019). Thus, further research is needed to examine how young children perceive and categorize more variable emotional stimuli (Ruba & Pollak, 2020), including how language relates to such perception and categorization.

One limitation of the study is that a relatively high number of 2-year-olds were excluded from both studies for failing the familiarization trials (N=14 per study). Although some participant data loss with young children is expected due to difficulty following instructions or engaging with the task, it is important to note that the present results from the 2-year-olds in particular may only generalize to those who are good at completing categorization tasks. It is also worth noting that our sample size was sufficiently powered to detect medium-large interaction effects. Thus, future studies may wish to determine whether smaller interaction effects may be detectable with larger sample sizes. Future studies may also want to randomly assign children to the earlier or later-learned emotion conditions. Children in this study were not randomly assigned to

participate in Study 1 (earlier-learned) or Study 2 (later-learned), so the cross-study analyses should be taken lightly, as random assignment to each study would be necessary to draw strong conclusions when comparing the two studies. Finally, it is worth bearing in mind that children here were tasked with sorting adult faces which had been previously normed by adult raters (Tottenham et al., 2009). Although this is common for research with children and infants (e.g., Kammermeier & Paulus, 2021; Lee et al., 2017), future research addressing the role of language in children's categorization of other children's faces (e.g., LoBue et al., 2017) would be of great value.

Lastly, a previous study with 2- to 7-year-old children (Study 1 of Russell & Widen, 2002), examined whether faces, labels, or the combination of faces and labels was most beneficial for helping children to determine if a face belonged in a "happy" or "angry" category. These researchers found that presenting children with emotion labels in conjunction with an exemplar face aided emotion categorization more than the face alone, but not more than a label alone across all the ages tested (including two- and three-year-olds). Russell and Widen's (2002) findings are in contrast to the results of Study 1 in the present paper, which found that labels did not improve categorization performance in two and three-year old children. Interestingly, both Russell and Widen's (2002) and the current study found that children did well when provided with the labels "happy" or "angry" and in the non-label condition when provided with "happy." However, children in the non-label condition performed considerably worse when queried on "angry" in the Russell and Widen (2002) study than they did in the current study. There were several methodological differences between the studies that may have affected the results. Children in the Russell and Widen (2002) study had to decide whether a face belonged in a single target category or should be left out of the category (e.g., "this box was for people who 'feel like this [face with angry expression] and only people who feel like this can go into the box. Does this person [pointing] feel the same as that one [pointing]?'). But more importantly, the distracters in the Russell and Widen (2002) study included a number of negative affect faces

that children frequently placed in the angry box, including disgust and fear. This lends support to the idea that labels may be more beneficial for helping children to discriminate between some of the later learned emotions, such as those in Study 2 of the current work. Taken together, the present results diverge from the findings of Russell and Widen (2002) regarding the role of labels for children's sorting of happy and angry emotion categories specifically but converge on the general conclusion that emotion words help children to sort later-learned emotion categories in a manner similar to categorizing other abstract content.

In sum, the current studies specifically examined whether labels influenced children's decisions regarding which emotion category, among multiple, a face belonged to. The current studies also looked at six different emotions to test the difference between earlier and later learned emotions, which is important for more comprehensively determining the role of language in emotion categorization as children develop and their understanding of emotions becomes more mature. Labels are helpful with this process of discriminating between categories to learn new words in other domains (Lupyan, et al., 2007), and the current studies show the same result for emotion categories, as labels were helpful for children's categorization of later learned emotions, but not earlier learned emotions. Taken together, the present studies offer unique insight into the role of language in emotion categorization. To conclude, the present studies suggest that labels may help children to form emotion categories beginning in infancy and through early childhood. In fact, at the end of our studies children in both conditions verbally identified the emotions in Study 1 with high accuracy but were substantially less accurate in Study 2. This may suggest that, early in development, labels help infants to form the earliest-learned emotion categories (e.g., happy, sad, angry), and that later in development labels continue to be helpful for learning more challenging or nuanced emotion categories (e.g., disgust, shame, annoyance). Addressing this potential developmental trend more directly will be important for future research. Altogether, the results of the present study provide additional support for theoretical perspectives that suggest the importance of labels for early childhood

emotion categorization.

References

Althaus, N., & Plunkett, K. (2016). Categorization in infancy: Labeling induces a persisting focus on commonalities. *Developmental Science, 19,* 770-780. doi:10.1111/desc.12358

Barker, M. S., Bidstrup, E. M., Robinson, G. A., & Nelson, N. L. (2020). “Grumpy” or “furious”? Arousal of emotion labels influences judgments of facial expressions. *PLoS ONE, 15*(7), 1–18. https://doi.org/10.1371/journal.pone.0235390

Baron-Cohen, S., Golan, O., Wheelwright, S., Granader, Y., & Hill, J. (2010). Emotion word comprehension from 4 to 17 years old: A developmental survey. *Frontiers in Evolutionary Neuroscience, 2*, 109. doi:10.3389/fnevo.2010.00109

Barrett, L. F. (2017). *How Emotions Are Made: The Secret Life of the Brain.* New York, NY: Houghton-Mifflin-Harcourt.

Barrett, L. F., Adolphs, R., Marsella, S., Martinez, A. M., & Pollak, S. D. (2019). Emotional expressions reconsidered: Challenges to inferring emotion from human facial movements. *Psychological Science in the Public Interest, 20*, 1-68. doi:10.1177/1529100619832930

Denham, S. A. (1986). Social cognition, prosocial behavior, and emotion in preschoolers: Contextual validation. *Child Development, 57,* 194-201. doi:10.1111/j.1467-8624.1986.tb00020.x

Ekman, P. (1992). An argument for basic emotions. *Cognition and Emotion, 6,* 169-200. doi:10.1080/02699939208411068

Ekman, P. (1999). Basic Emotions. In T. Dalgleish & M. Power (Eds.) *Handbook of Cognition and Emotion* (pp. 45-60). Sussex, U.K.: John Wiley & Sons, Ltd.

Ekman, P., & Cordaro, D. (2011). What is meant by calling emotions basic. *Emotion Review, 3,* 364-370. doi:10.1177/1754073911410740

Erickson, L. C., Thiessen, E. D., & Estes, K. G. (2014). Statistically coherent labels facilitate categorization in 8-month-olds. *Journal of Memory and Language, 72,* 49-58. doi:10.1016/j.jml.2014.01.002

Faul, F., Erdfelder, E., Buchner, A., & Lang, A. G. (2009). Statistical power analyses using G*Power 3.1: Tests for correlation and regression analyses. *Behavior Research Methods, 41*, 1149-1160. doi:10.3758/BRM.41.4.1149

Fugate, J. M., Gouzoules, H., & Barrett, L. F. (2010). Reading chimpanzee faces: Evidence for the role of verbal labels in categorical perception of emotion. *Emotion, 10*, 544-554. doi:10.1037/a0019017

Fulkerson, A. L., & Haaf, R. A. (2003). The influence of labels, non-labeling sounds, and source of auditory input on 9- and 15-month-olds' object categorization. *Infancy, 4,* 349-369. doi:10.1207/S15327078IN0403_03

Gentner, D., Anggoro, F. K., & Klibanoff, R. S. (2011). Structure mapping and relational language support children's learning of relational categories. *Child Development, 82,* 1173-1188. doi:10.1111/j.1467-8624.2011.01599.x

Gentner, D., & Markman, A. B. (1995). Similarity is like analogy: Structural alignment in comparison. In C. Cacciari (Ed.), *Similarity in language, thought and perception* (pp. 111-147). Brussels: BREPOLS.

Gentner, D., & Namy, L. L. (1999). Comparison in the development of categories. *Cognitive Development, 14,* 487-513. doi:10.1016/S0885-2014(99)00016-7

Gliozzi, V., Mayor, J., Hu, J. F., & Plunkett, K. (2009). Labels as features (not names) for infant categorization: A neurocomputational approach. *Cognitive Science, 33,* 709-738. doi:10.1111/j.1551-6709.2009.01026.x

Grosse, G., Streubel, B., Gunzenhauser, C., & Saalbach, H. (2021). Let's talk about emotions: The development of children's emotion vocabulary from 4 to 11 years of age. *Affective Science, 2*, 150-162. doi:10.1007/s42761-021-00040-2

Hoemann, K., Gendron, M., & Barrett, L. F. (2017). Mixed emotions in the predictive brain. *Current Opinions in Behavioral Sciences, 15*, 51-57. doi:10.1016/j.cobeha.2017.05.013

Hoemann, K., Xu, F., & Barrett, L. F. (2019). Emotion words, emotion concepts, and emotional development in children: A constructionist hypothesis. *Developmental Psychology, 55*, 1830-1849. doi:10.1037/dev0000686

Izard, C. E., Fine, S., Schultz, D., Mostow, A. J., Ackerman, B., & Youngstrom, E. (2001). Emotion knowledge as a predictor of social behavior and academic competence in children at risk. *Psychological Science, 12,* 18-23. doi:10.1111/1467-9280.00304

JASP Team (2020). JASP (Version 0.8.6)[Computer software].

Jeffreys, H. (1961). Theory of probability (3rd Ed.). Oxford, UK: Oxford University Press.

Johanson, M., & Papafragou, A. (2016). The influence of labels and facts on children's and adults' categorization. *Journal of Experimental Child Psychology, 144,* 130-151. doi:10.1016/j.jecp.2015.11.010

Kammermeier, M., & Paulus, M. (2021). Maternal sensitivity and non-intrusiveness at 12 months predict attention to emotional facial expressions at 24 months: A cross-lagged panel approach. *Social Development,* online ahead of print. doi:10.1111/sode.12561

Lakens, D., McLatchie, N., Isager, P. M., Scheel, A. M., & Dienes, Z. (2020). Improving inferences about null effects with Bayes Factors and Equivalence Tests. *The Journals of Gerontology: Series B, 75*, 45-57. doi:10.1093/geronb/gby065

Lee, V., Besada, M., & Rutherford, M. D. (2017). Individual differences in emotional expression discrimination are associated with emotion label production in toddlers. *European Journal of Developmental Psychology, 15*, 506-516. doi:10.1080/17405629.2017.1314214

Lindquist, K. A., & Gendron, M. (2013). What's in a word? Language constructs emotion perception. *Emotion Review, 5,* 66-71. doi:10.1177/1754073912451351

Lindquist, K. A., MacCormack, J. K., & Shablack, H. (2015). The role of language in emotion: Predictions from psychological constructionism. *Frontiers in Psychology, 6,* Article 444. doi:10.3389/fpsyg.2015.0044

LoBue, V., Baker, L., & Thrasher, C. (2017). Through the eyes of a child: Preschoolers' identification of emotional expressions from the child affective facial expression (CAFE) set. *Cognition and Emotion, 32*, 1122-1130. doi:10.1080/02699931.2017.1365046

Loewenstein, J., & Gentner, J. (2005). Relational language and the development of relational mapping. *Cognitive Psychology, 50*, 315-353. doi:10.1016/j.cogpsych.2004.09.004

Lupyan, G., Rakison, D. H., & McClelland, J. L. (2007). Language is not just for talking: Redundant labels facilitate learning of novel categories. *Psychological Science, 18,* 1077-1083. doi:10.1111/j.1467-9280.2007.02028.x

Nelson, N. L., & Russell, J. A. (2016a). Building emotion categories: Children use a process of elimination when they encounter novel expressions. *Journal of Experimental Child Psychology*, *151*, 120–130. https://doi.org/10.1016/j.jecp.2016.02.012

Nelson, N. L., & Russell, J. A. (2016b). A facial expression of pax: Assessing children's "recognition" of emotion from faces. *Journal of Experimental Child Psychology*, *141*, 49–64. https://doi.org/10.1016/j.jecp.2015.07.016

Nelson, N. L., Nowicki, E., Diemer, C. M., Sangster, K., Cheng, C., & Russell, J. A. (2018). Children can create a new emotion category through a process of elimination. *Cognitive Development, 47*, 117–123. https://doi.org/10.1016/j.cogdev.2018.04.006

Nook, E. C., Sasse, S. F., Lambert, H. K., McLaughlin, K. A., & Somerville, L. H. (2017). Increasing verbal knowledge mediates development of multidimensional emotion representations. *Nature Human Behaviour, 1*, 881-889. doi:10.1038/s41562-017-0238-7

Nook, E. C., Stavish, C. M., Sasse, S. F., Lambert, H. K., Mair, P., McLaughlin, K. A., & Sommerville, L. H. (2020). Charting the development of emotion comprehension and abstraction from childhood to adulthood using observer-rated and linguistic measures. *Emotion, 20*, 773-792. doi:10.1037/emo0000609

Oakes, L. M., & Rakison, D. (2019). *Developmental Cascades: Building the Infant Mind.* New York, NY: Oxford University Press.

Olson, D. R., Astington, J. W., & Harris, P. L. (1988). Introduction. In J. W. Astington, P. L. Harris, & D. R. Olson (Eds.), *Developing theories of mind* (pp. 1-8). Cambridge, England. Cambridge University Press.

Perugini, M., Gallucci, M., & Costantini, G. (2018). A Practical Primer To Power Analysis for Simple Experimental Designs. *International Review of Social Psychology, 31*, 20. doi:10.5334/irsp.181

Plunkett, K., Hu, J. F., & Cohen, L. B. (2008). Labels can override perceptual categories in early infancy. *Cognition, 106,* 665-681. doi:10.1016/j.cognition.2007.04.003

Reschke, P. J., Walle, E. A., & Dukes, D. (2017). Interpersonal development in infancy: The interconnectedness of emotion understanding and social cognition. *Child Development Perspectives*, *11*, 178-183. doi:10.1111/cdep.12230

Rhodes, M., Leslie, S. J., Bianchi, L., & Chalik, L. (2018). The role of generic language in the early development of social categorization. *Child Development, 89*, 148-155. doi:10.1111/cdev.12714

Ridgeway, D., Waters, E., & Kuczaj, S. A. (1985). Acquisition of emotion-descriptive language: Receptive and productive vocabulary norms for ages 18 months to 6 years. *Developmental Psychology, 21,* 901-908. doi:10.1037/0012-1649.21.5.901

Ruba, A. L., Harris, L. T., & Wilbourn, M. P. (2020). Examining preverbal infants' ability to map labels to facial configurations. *Affective Science, 2,* 142-149. doi:0.1007/s42761-020-00015-9

Ruba, A. L., Meltzoff, A. N., & Repacholi, B. M. (2020). Superordinate categorization of negative facial expressions in infancy: The influence of labels. *Developmental Psychology, 56*, 671-685. doi:10.1037/dev0000892

Ruba, A. L., & Pollak, S. (2020). The development of emotion reasoning in infancy and early childhood. *Annual Review of Developmental Psychology, 2*, 503-531. doi:10.1146/annurev-devpsych-060320-102556

Ruba, A. L., & Repacholi, B. M. (2019). Do preverbal infants understand discrete facial expressions of emotion? *Emotion Review, 12*, 235-250. doi:10.1177/1754073919871098

Ruba, A. L., Wilbourn, M. P., Ulrich, D. M., & Harris, L. T. (2018). Constructing emotion categorization: Insights from developmental psychology applied to a young adult sample. *Emotion, 18,* 1043-1051. doi:10.1037/emo0000364

Russell, J. A., & Widen, S. C. (2002). A label superiority effect in children's categorization of facial expressions. *Social Development, 11,* 30-52. doi:10.1111/1467-9507.00185

Schwartz, G. M., Izard, C. E., & Ansul, S. E. (1985). The 5-month-old's ability to discriminate facial expressions of emotion. *Infant Behavior and Development, 8*, 65-77. doi:10.1016/S0163-6383(85)80017-5

Shablack, H., Stein, A. G., & Lindquist, K. A. (2020). Comment: A role of language in infant emotion concept acquisition. *Emotion Review*, *12*(4), 251-253. doi:10.1177/1754073919897297

Sloutsky, V. M., & Fisher, A. V. (2004). When development and learning decrease memory: Evidence against category-based induction in children. *Psychological Science, 15*(8), 553–558. https://doi.org/10.1111/j.0956-7976.2004.00718.x

Tottenham, N., Tanaka, J., Leon, A. C., McCarry, T., Nurse, M., Hare, T. A., Marcus, D. J., Westerlund, A., Casey, B. J., & Nelson, C. A. (2009). The NimStim set of facial expressions: Judgments from untrained research participants. *Psychiatry Research, 168*(3), 242-294. doi:10.1016/j.psychres.2008.05.006

Voltmer, K., & Von Salisch, M. (2017). Three meta-analyses of children's emotion knowledge and their school success. *Learning and Individual Differences, 59,* 107-118. doi:10.1016/j.lindif.2017.08.006

Walker-Andrews, A. S., & Grolnick, W. (1983). Discrimination of vocal expressions by young infants. *Infant Behavior and Development, 6*, 491-498. doi:10.1016/S0163-6383(83)90331-4

Waxman, S. R., & Markow, D. B. (1995). Words as invitations to form categories: Evidence from 12- to 13-month-old infants. *Cognitive Psychology, 29*(3), 257–302. https://doi.org/10.1006/cogp.1995.1016

Widen, S. C. (2013). Children's interpretation of facial expressions: The long path from valence-based to specific discrete categories. *Emotion Review, 5*(1), 72–77. https://doi.org/10.1177/1754073912451492

Widen, S. C., & Russell, J. A. (2003). A closer look at preschoolers' freely produced labels for facial expressions. *Developmental Psychology, 39,* 114-128. doi:10.1037/0012-1649.39.1.114

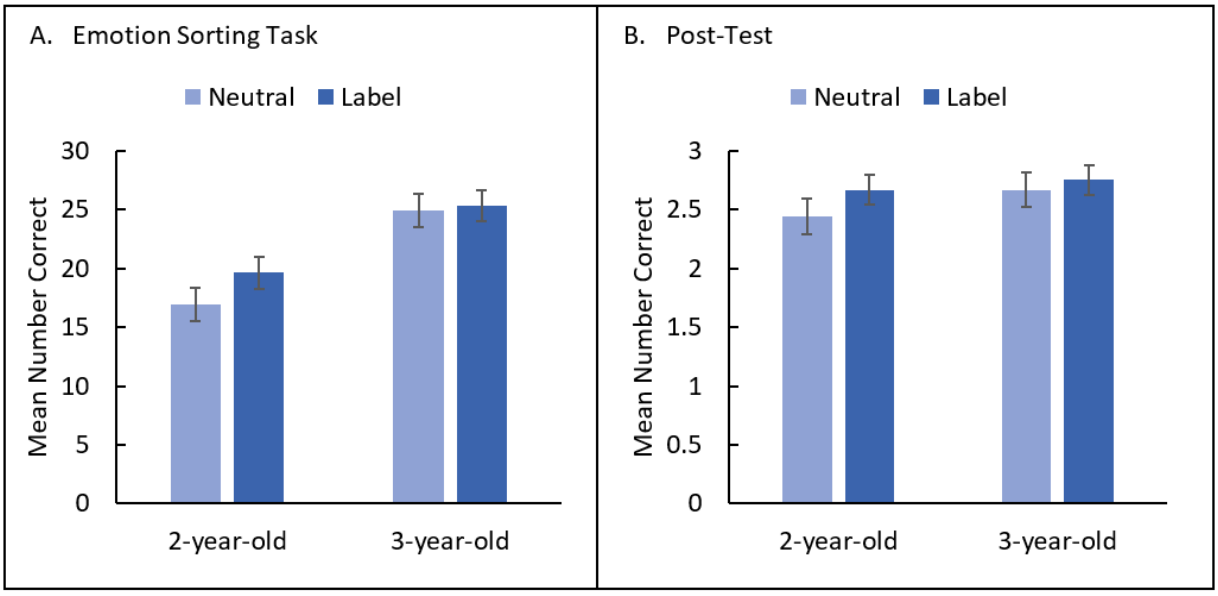
Widen, S. C., & Russell, J. A. (2008a). Children acquire emotion categories gradually. *Cognitive Development, 23,* 291-312. doi:10.1016/j.cogdev.2008.01.002

Widen, S. C., & Russell, J. A. (2008b). Young children's understanding of other's emotions. In Lewis, M., Haviland-Jones, J. M., & Barrett, L. F. (Eds.) *Handbook of Emotions* (3rd edition) (pp. 348-363). New York, NY. The Guilford Press.

Wu, Y., Muentener, P., & Schulz, L. E. (2017). One- to four-year-olds connect diverse positive emotional vocalizations to their probable causes. *Proceedings of the National Academy of Sciences, 114*, 11896-11901. doi:/10.1073/pnas.170771511

Figure 1

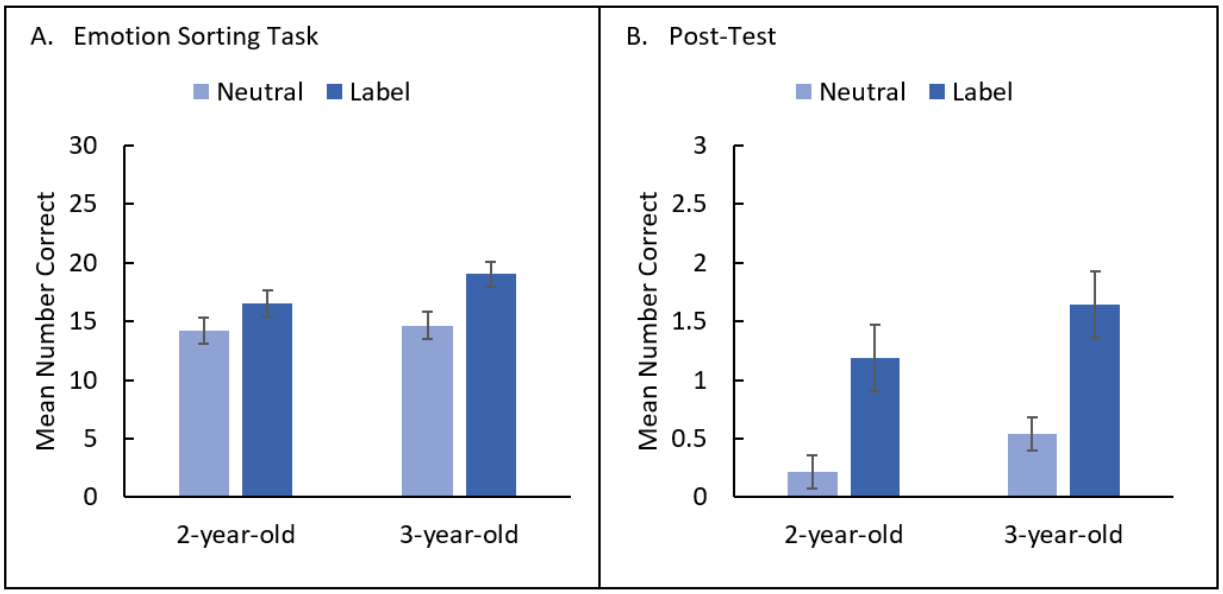
Bar plots depicting (Panel A) the average correct choices in sorting emotion faces and (Panel B) the average correct choices in the post-test in each experimental condition, separated by age group.



Note. These data are from Study 1, and error bars indicate standard error.

Figure 2.

Bar plot depicting (Panel A) the average correct choices in sorting emotion faces and (Panel B) the average correct choices in the post-test in each experimental condition, separated by age group.



Note. These data are from Study 2, and error bars indicate standard error.