How Do Emotion Words Impact Children's Emotion Learning?

Marissa Ogren, Vanessa LoBue, Catherine M. Sandhofer

Funded by:

- Eunice Kennedy Shriver National Institute of Child Health and Human Development
- James S. McDonnell Foundation

How do Emotion Words Impact Children's Emotion Learning?

Marissa Ogren[1]

Vanessa LoBue[1]

Catherine M. Sandhofer[2]

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

Correspondence

Correspondence concerning this article should be addressed to Marissa Ogren,

101 Warren St Room B21, Newark, NJ 07102

Contact: Marissa.ogren@rutgers.edu, (973)-353-3938

The authors have no conflicts of interest to disclose

Acknowledgments

This work was supported by NIH grant F32-HD105316 and the James McDonnell Foundation Scholar Award for Understanding Human Cognition. The authors would like to thank the parent and child volunteers, Lauren Huang and Idella Smolyar, and the members of the Rutgers University Child Study Center.

Author Note

These results have been previously presented in poster format at the Cognitive Society Conference which took place July 26-29, 2023 in Sydney, Australia. Study materials (vignettes and images) are available in supplemental materials, and data are available upon request.

OSF Repository Link: https://osf.io/c3reh/

CRediT Statement

MO: Conceptualization, Formal Analysis, Funding Acquisition, Investigation, Methodology, Project Administration, Supervision, Visualization, Writing- Original Draft

VL: Funding acquisition, Resources, Writing- Reviewing and Editing

CMS: Conceptualization, Methodology, Resources, Writing- Reviewing and Editing

Date Submitted: 9/2/2023

Revision Submitted: 2/6/2024

Abstract

Previous research suggests that the use of emotion labels helps children to learn about emotions. However, the mechanism behind this relation remains somewhat elusive. The present study examined 3-year-old children's (*N*=72; M_{age}=3.51 years; 42 female) ability to match faces to emotional vignettes, and the role that the use of emotion labels plays in this process. Parents identified participating children as White (*N*=37), Multiracial (*N*=17), African-American/Black (*N*=5), Asian (*N*=5), Hispanic (*N*=3), Latino (*N*=2), South Asian/Indian (*N*=1), Middle Eastern (*N*=1), and Other (*N*=1), and most children had a parent with a college degree (*N*=66). After a pre-test, children heard either explicit emotion labels ("she feels *annoyed*"), novel labels ("she feels *wuggy*"), or irrelevant information ("she sits down") paired with a vignette and associated facial configuration. Children were then tested again at post-test for evidence of learning. Results revealed that children only improved from pre- to post-test in the explicit label condition, demonstrating that explicit emotion labels, which are likely to be familiar to children, facilitate children's learning of emotion information. Altogether, our results suggest that familiarity with emotion words from prior daily experience may best explain how emotion words influence children's learning about emotions.

Keywords: Emotion, Learning, Labels, Early Childhood

Public Significance Statement

Hearing emotion words helps young children to learn about emotion categories. However, we do not yet know exactly *how*. This study finds that children learn best from emotion words that they have likely heard before over a period of years, suggesting that there is a benefit to hearing the same emotion words many times throughout their daily experiences.

How Do Emotion Words Impact Children's Emotion Learning?

Learning about emotions is a crucial skill for development. Children who talk more about emotions are rated as more likable by their peers (Fabes et al., 2001), and children with higher levels of emotion understanding tend to have greater academic success (Voltmer & von Salisch, 2017). Further, young children who have a better understanding of others' emotions tend to demonstrate more social competence and better peer relationships (Denham et al., 2003), likely because the ability to make inferences about others' emotions helps to provide insight into others' goals and behaviors (Reschke et al., 2017). Thus, developing an understanding of emotions holds long-term implications for children's social-cognitive development. For this reason, it is crucial to better understand how children learn about emotions early in life, and what factors may account for individual differences in children's early emotion understanding.

Previous theoretical work has suggested that access to emotion words may be one factor that influences how individuals think and learn about emotions (Barrett, 2017; Hoemann et al., 2019). Some research with adults supports this perspective, indicating that when adults do not have access to emotion words, they show disrupted processing and sorting of emotional information (e.g., Gendron et al., 2012). Further, recent research with young children demonstrates that hearing emotion words helps children learn associations between emotion-eliciting events and facial configurations (Ogren & Sandhofer, 2022).

But why exactly do words help children learn emotion concepts? Because emotions are abstract concepts, inferring emotional meaning involves aggregating information across a complex array of faces, vocalizations, body movements, and contextual information that are all dynamically changing over time (Hoemann et al., 2017). In this way, emotion words may help children learn emotion categories in the same way that words help children to develop an understanding of other types of abstract or complex categories (Hoemann et al., 2020)—namely by helping individuals to identify the similarities among variable emotional events that are meaningful for categorization. For example, although people differ widely in how they express

anger, using a label (“anger”) to describe variable exemplars of this emotion may help draw children's attention to the important similarities across instances of anger.

Further, reasoning about abstract categories makes notable demands on working memory and attention (Hummel & Holyoak, 1997). Indeed, preschool children have difficulty learning about abstract concepts in the presence of other salient features. However, using labels helps (Gentner & Rattermann, 1991), likely by allowing children to identify what abstract aspects of variation are meaningful for category formation, and thereby making transfer to new situations easier as well (Son et al., 2010). Emotion words may operate similarly, by helping children to link abstract emotion categories to a predictable, abstract set of features. What remains unclear, however, is precisely how emotion labels benefit emotion learning. Do labels simply serve as “tags” or connections to help children draw comparisons, and thus any novel word presented consistently across trials could help to facilitate in-the-moment categorization? Or does the label need to be familiar from children’s prior experience, whereby some level of familiarity with the semantic content helps children to compare across real-world experiences?

Prior work has shown that within a brief experimental paradigm, emotion words (e.g., “nervous”) but not broader terms (e.g., “feels bad”) help young children learn about emotions (Ogren & Sandhofer, 2022). In one study, 3-year-old children selected which face matched multiple scenarios at pre-test and post-test. Between pre- and post-test children heard labels (e.g., “nervous” or “feels bad”) after hearing a scenario and seeing the stereotypical face matching the scenario. Only children who heard specific emotion labels (e.g., “nervous”) but not broader terms (e.g., “feels bad”) improved from pre- to post-test. This may be because the emotion labels were sufficiently specific to facilitate learning the association between faces and scenarios, but the broader terms were too generic and often specific described something other than emotion in children’s prior experience (e.g., “the milk has gone *bad*” or “they made a *bad* decision”). However, it remains an open question whether the emotion words in the prior study were helpful because children had previously heard them used in emotional situations in the

real world, or whether simply hearing *any* specific labels helps children connect various instances of a single emotion. In non-emotional domains, simply hearing the same label with multiple stimuli can "invite" children to draw comparisons (Brown, 1958). Labels may also induce abstraction (Gentner & Namy, 1999), helping children to expand beyond the specific features in front of them to form more abstract concepts. Thus, it may be that simply presenting children with *any* specific labels, regardless of whether they are familiar or completely novel, may facilitate the necessary comparison and abstraction to learn about emotions.

In contrast, it may be that prior experience with actual emotion words in the real world is what facilitated children's improved performance. That is, even though children may not have a firm grasp of these emotion words at first, drawing on their prior experience with them may help learning within the task. Prior research has shown that children's experience with familiar words can impact how they learn new abstract concepts. For example, Rattermann and Gentner (1998) found that using familiar terms like "mommy", "daddy", and "baby" helped children to learn size relations. This may be because familiar words allow children to draw on their rich past experience to evoke schemas that are flexible enough to be applied to new situations, and these schemas may then facilitate learning about new abstract concepts (Son et al., 2012). Teasing apart the potential benefits of novel labels versus actual emotion words that may be familiar from children's daily experience is important for understanding how labels may help emotion learning.

The present study aimed to examine the role of familiar emotion words in aiding emotion category learning. To do so, we closely mirrored the method of Ogren and Sandhofer (2022), who reported that explicit emotion labels help 3-year-old children to learn about emotions. Here, we tasked 3-year-old children with identifying which of a selection of faces best matched how a character would feel in various emotional situations. Children participated in learning trials, where they were randomly assigned to hear either explicit emotion labels (e.g., "she feels annoyed"), irrelevant information (e.g., "she sits down"), or novel labels (e.g., "she feels

wuggy"). The Explicit Label and Irrelevant Information conditions were the same as those used in the original Ogren and Sandhofer (2022) study, which allowed us to determine whether the previous results replicated. The Novel Label condition extended this prior research and allowed us to test whether hearing any label would (similar to the Explicit Condition) or would not (similar to the Irrelevant Condition) lead to improved performance.

Method

Participants

Seventy-two children (42 female) participated in this study, ranging in age from 3.04 to 3.95 years (M_{age}=3.51 years, SD_{age}=0.26 years). Data were collected between May 2022 and November 2022. Participants were recruited from across the United States, with recruitment occurring from two different university databases (UCLA and Rutgers), through Children Helping Science, and through social media posts. The final sample included children from the following racial/ethnic backgrounds: White (*N*=37), Multiracial (*N*=17), African-American/Black (*N*=5), Asian (*N*=5), Hispanic (*N*=3), Latino (*N*=2), South Asian/Indian (*N*=1), Middle Eastern (*N*=1), Other (*N*=1). Sixty-six of the participants had a parent who had graduated from college. This study received approval from the UCLA and Rutgers University Institutional Review Boards (IRB# 21-001623 and Pro2021002363, respectively) and complied with APA ethical standards. Parents provided informed consent and were asked to report the number of emotion words their child says from a list provided (happy, sad, angry, mad, afraid, scared, annoyed, nervous, disgusted). This list contained emotion words relevant to the present study, including our target emotion words (e.g., afraid, angry, annoyed, disgusted), synonyms for some of those words (e.g., scared, mad), and additional emotion words that served as distractors in the task (e.g., happy, sad). All participants received a $5 gift card.

Transparency and Openness

Data from 8 additional children were collected but excluded for selecting the same face for all trials (*N*=3), not completing all trials (*N*=3), or requiring parental translation (*N*=2). The

sample size was determined using a G*power 3.1 power analysis (Faul et al., 2009). Specifically, we used the reported effect size comparing explicit and irrelevant trials from Ogren and Sandhofer (2022) of d=0.75, an alpha level of 0.05, and power of 0.8 to determine that the necessary sample size for our proposed ANOVA comparing three groups would be 72 (24 children per group). Study design and hypotheses were pre-registered (https://osf.io/2w3mg). Recruitment continued until the target, pre-registered sample size of 72 participants was achieved.

Materials

The present study included four target emotions (afraid, angry, annoyed, disgusted). According to Ogren and Sandhofer (2021), these are all emotion labels that 1- to 3-year-old children hear in speech from mothers and are therefore likely to be somewhat familiar (although frequencies among these emotion labels vary). Further, prior work suggests that children tend not to fully label and differentiate the emotion categories of anger and fear until close to 48 months of age (Widen & Russell, 2010), suggesting that 3-year-old children would likely be well-suited to learn about these 4 emotions in the present study. Stimuli included images of facial configurations commonly associated with each emotion, brief vignettes, and drawings associated with the vignettes. This closely mirrored the Ogren and Sandhofer (2022) study materials with some modification, as the original study examined the target emotions of annoyed, disgusted, and nervous. Including new emotions allowed us to examine whether the prior findings generalized beyond the original set of emotion categories.

Twenty-four vignettes conveyed angry, annoying, fearful, and disgusting scenarios (6 per emotion; see Supplemental Material). Vignettes for annoyed and disgusted were drawn directly from the original Ogren and Sandhofer (2022) study, while new vignettes were created for fear and anger. Each vignette included a drawing that provided no clues to the emotion but could help children remember the story.

For faces, 11 undergraduate women were recorded while they were asked to think about a time when they felt a target emotion (anger, annoyance, disgust, fear, happiness, and sadness) and to convey that emotion through their face. Only images of women were used, as women tend to contribute more to childcare for young children (Fillo et al., 2015), and thus these faces were more likely to be familiar to our 3-year-old participants. This decision also mirrored the stimuli used in the original Ogren and Sandhofer (2022) study. Faces for anger, annoyance, disgust, and fear were included as both target and distractor choices (depending on the specific trial), and faces for happiness and sadness were included as distractors/incorrect choices.

Stimulus Validation

To determine whether adults interpreted these faces as belonging to the intended emotion category, twenty-six adults (3 male) were surveyed. Each adult was asked to select which emotion label (afraid, angry, annoyed, disgusted, happy, nervous, or sad) best described each of the 66 images (11 women, 6 images per woman). Based on these responses, the examples of each emotion category with the highest agreement were chosen. On average, the adult raters agreed on the intended emotion 84.4% of the time (Afraid=75.5%, Angry=77.4%, Annoyed=75.2%, Disgusted=93.6%, Happy=95.4%, Sad=89.2%). Raters agreed on the intended emotion of these images at well above chance levels.

We also surveyed eleven additional adults (1 male) to determine whether adults interpreted the vignettes in isolation as belonging to the intended emotion category. Each adult read the twenty-four vignettes in written form and was asked to indicate which emotion label (afraid, angry, annoyed, disgusted, happy, or sad) best matched how the character in each story would be feeling. On average, the adult raters agreed on the intended emotion 90.2% of the time (Afraid=91%, Angry=71.2%, Annoyed=98.5%, Disgusted=100%). Thus, raters also agreed on the intended emotion of the vignettes at well above chance levels.

Procedure

The procedure assessed children's emotion category learning in a live task over Zoom. All children participated with a parent present. Prior to the start of the experiment, children were randomly assigned to either the Explicit Label, Novel Label, or Irrelevant Information condition. All stimuli were presented through screen sharing over Zoom. Parents were asked to stand behind their child and not to interfere with their child's choices, but were asked to report which option their child selected if the child's point was unclear to the experimenter over the Zoom call. Children in this study participated in 8 pre-test trials, 8 learning trials, and 8 post-test trials. Four separate study orders were created to randomize which vignettes occurred during pre-test, learning, and post-test with the constraint that 2 vignettes of each target emotion (angry, afraid, annoyed, and disgusted) needed to be presented for each trial type.

Pre-Test. For all trials, the experimenter began by reading the brief vignette to the child. While the vignette was read, the drawing associated with that vignette was shown on the screen to help children remember the story's content. All experimenters were trained to deliver the vignettes consistently, without changes to tone of voice or any non-verbal vocalizations. Then children were asked to select which face (from four options) the character would feel. Of the four options, one was always the correct choice (e.g., an annoyed face after an annoyed vignette). One of the remaining three options was a different target emotion category from the study (e.g., angry, afraid, or disgusted). The remaining two faces were always happy and sad, representing one positive and one negative valence distractor option. All four faces presented on a given trial were always of the same woman. The experimenter recorded children's responses to all 8 trials, then advanced to the learning trials.

Learning. During learning trials, children saw only the drawing associated with a vignette and one face from the same emotion category while the experimenter read the relevant vignette. After reading the vignette, the experimenter provided different information depending on the study condition. If the child was in the Explicit Label condition, they heard the experimenter label the specific emotion category (e.g., "she feels *annoyed*"). In the Novel Label

condition, they heard the experimenter provide a novel label for the character's emotion (e.g., "she feels *wuggy*"). In the Irrelevant Information condition, they heard the experimenter provide information irrelevant to emotion (e.g., "she sneezes"). Irrelevant information was included as a control condition to determine whether simply seeing one face paired with one vignette during learning trials was sufficient to improve child performance.

For all conditions, the information paired with the emotion categories was consistent (e.g., disgusted faces and vignettes were always paired with either "she feels disgusted", "she feels blickety" or "she sits down"). In each condition children were then asked to point to the person (e.g., "Can you point to the girl who feels *blickety*?") and then asked to verbally provide the information back to the experimenter (e.g., "and how does she feel?"/"and what did she do?"). This provided children with additional opportunities to engage with the information that was provided verbally, and potentially help them better encode the learning trial information. Immediately after the 8th learning trial, children advanced to the post-test trials.

Post-Test. These trials were structured identically to the pre-test trials but with different vignettes. None of the vignettes repeated across pre-test, learning, or post-test trials, necessitating generalization of any learned information to new vignettes. Although some of the faces did repeat from pre-test to post-test (but never from learning trials), the same specific face was never the correct option at both pre-test and post-test. For example, if Woman A's fearful face was the correct choice at pre-test, a different image of Woman A was the correct choice at post-test (either angry, annoyed, or disgusted). Study vignettes and images are available in supplemental materials, and data are available upon request. Examples of possible test and learning trials are presented in Figures 1 and 2, respectively.

Results

Descriptive Statistics

Given that there were 4 possible response options and 8 vignettes at pre-test, if children responded at chance levels, children should answer correctly on 2 trials. The average number

of correct responses at pre-test was 2.79 (*SD*=1.23; range 0-6). No child answered all questions correctly at pre-test. Overall, these descriptive statistics suggest that our task was challenging for the 3-year-old children . This was precisely as intended, to ensure that there was no ceiling effect at pre-test and to allow children an opportunity to learn and improve at post-test. Importantly, there also were no differences between children's performance at pre-test by condition ($F(2,69)=0.32$, $p=.725$, $\eta_p^2=0.01$). For parent reports of children's emotion vocabulary, on average parents reported that their child produced 6.17 of the 9 assessed emotion words (*SD*=1.69), with values ranging from 1 to 9.

Effect of Condition

To examine our primary research question, we conducted a one-way ANOVA to test whether children's learning from pre-test to post-test differed by condition. Results revealed that children's difference score (post-test score minus pre-test score) did significantly differ by condition ($F(2,69)=4.51$, $p=.014$, $\eta_p^2=0.12$) (Explicit Label=0.96,*SD*=1.57; Novel Label=0.08, *SD*=1.69; Irrelevant Information=-0.38,*SD*=1.41) (see Figure 3). We then conducted three follow-up t-tests to directly compare each pair of conditions. Results revealed that the difference score for the Explicit Label condition was significantly higher than the Irrelevant Information condition ($t(46)=3.09$, $p=.003$, $d=0.89$), and that the Explicit Label condition trended toward higher performance than the Novel Label condition ($t(46)=1.86$, $p=.070$, $d=0.54$). The Novel Label and Irrelevant Information conditions did not significantly differ ($t(46)=1.02$, $p=.313$, $d=0.29$). Further, additional post-hoc analyses indicated that only the difference score for the Explicit Label condition significantly differed from chance performance of 0, ($t(23)=2.98$, $p=.007$, $d=0.61$), while the Novel Label ($t(23)=0.24$, $p=.811$, $d=0.05$) and Irrelevant ($t(23)=-1.30$, $p=.205$, $d=0.27$) conditions did not. An emotion-by-condition interaction was not examined, as such an analysis would have limited power due to each child completing only 2 trials per emotion at both pre and post-test.

To supplement the primary analyses, we conducted post-hoc Bayesian analyses using JASP (JASP Team, 2020). These analyses provide an informative complement to the previously reported frequentist statistics, as they can inform not only whether a result is significant, but also how much evidence there is for or against the null hypothesis (Lakens et al., 2020). Results revealed moderate evidence for the effect of condition on learning (BF_{10}=3.49). Further, when directly comparing the conditions to each other, there was strong evidence for higher performance in the Explicit condition compared to the Irrelevant condition (BF_{10}=11.54), anecdotal evidence for a difference between conditions when comparing the Explicit condition to the Novel condition (BF_{10}=1.14), and anecdotal evidence for the null hypothesis when comparing the Irrelevant condition to the Novel condition (BF_{10}=0.43). Thus, the null results in the frequentist analyses should be interpreted with caution. Finally, when directly comparing the difference score in each condition to 0, we found strong evidence for above-chance difference scores in the Explicit condition (BF_{10}=6.84), anecdotal evidence for the null hypothesis in the Irrelevant condition (BF_{10}=0.46), and moderate evidence for the null hypothesis in the Novel condition (BF_{10}=0.22), suggesting caution in interpreting whether or not there was a deviation from 0 in the Irrelevant condition, but support for the interpretation that children did learn in the Explicit condition, and no difference from pre- to post-test in the Novel condition.

Emotion Vocabulary

We further examined whether the effect of condition was mediated by other variables, such as the number of emotion words in the child's vocabulary. To examine this, we included the number of emotion words parents reported that their child said (out of 9 possible) along with study condition and the emotion vocabulary*condition interaction in a regression model predicting children's difference score. We found that neither emotion vocabulary independently (β=-.02, p=.936) nor the interaction between emotion vocabulary and condition (β=.21, p=.646) significantly predicted children's difference score, suggesting that children's emotion vocabulary prior to the experiment did not influence their ability to learn about emotions within the study.

However, given that parents were only asked to report on 9 emotion words and most children produced more than half of these words, it is possible that there was not sufficient variability to detect an interaction with emotion vocabulary. Thus, this null result should be interpreted with caution, and future research may wish to account for the base rates of each word in the English language.

Child Age

Next, in a similar regression model, we examined whether child age related to children's emotion learning in the present study. The results revealed that neither child age (β=.08, p=.668) nor the interaction between child age and study condition (β=.96, p=.566) significantly predicted children's difference score.

Child Gender

Finally, we tested whether child gender predicted children's emotion learning. Neither child gender independently (β=-.02, p=.908) nor the interaction between child gender and condition (β=.26, p=.206) significantly predicted children's difference score. Taken together, these follow-up analyses indicate that children's learning of the emotion categories was not significantly influenced by the child's emotion vocabulary, age, or gender, suggesting that this pattern of emotion learning may be relatively stable across such individual difference factors.

Discussion

The present study examined whether explicit labels, novel labels, or irrelevant information would differentially impact how well children learned the relations between facial configurations and their corresponding emotional vignettes. Children's performance significantly increased from pre- to post-test only in the Explicit Label condition. This demonstrates that children improved after hearing the explicit emotion label, but not the completely novel label. Further, performance in the Explicit Label condition was significantly higher than in the Irrelevant Information condition and trended toward better performance than in the Novel Label condition. The Novel Label condition, however, did not significantly differ from the Irrelevant Information

condition, and the novel labels did not appear to either help or hinder children's performance, with an average difference score in this condition that was very close to 0, and Bayesian analyses indicating moderate evidence for the null effect. Further, children in the Irrelevant condition had a negative difference score, suggesting potentially worse performance at post-test, although this one-sample t-test was not significant. This may suggest that the irrelevant information in this control condition was potentially distracting for some children, although any hindrance effect did not lead to a difference score that differed significantly from 0. This pattern of results has two important implications. First, it aligns with prior research indicating that labels can facilitate children's learning of emotion categories. Second, and unique to this study, we found that using explicit emotion labels—which children have likely heard in the past—facilitates learning, while using completely novel labels may not.

These results are consistent with work in the category learning literature, which points to the importance of labels for highlighting similarities across category instances (e.g., Sloutsky, 2003) and forming representations of abstract categories (Loewenstein & Gentner, 2005). Aligning with this literature, the present results demonstrate that emotion words helped children to learn emotion categories. This study provides an important contribution to the existing literature by replicating previous work reporting that children learned more about emotions when given Explicit Labels compared to Irrelevant Information, and does so using additional emotion categories, with recruitment from two different geographic regions as well as online, and with a larger sample size. Thus, this replication provides further confidence that emotion labels help children to learn about emotions, at least insofar as they help children associate facial configurations with emotional situations.

Further, the current findings add to the literature by suggesting that for abstract emotion learning in particular, it is likely that experiences with explicit emotional labels in the real world—and not just hearing new words in a laboratory setting—are crucial for facilitating this learning (c.f., Rattermann & Gentner, 1998). However, it is important to keep in mind that *even explicit*

emotion labels are completely novel for children at some point in development. Therefore, it is likely that having heard these words in relevant emotional contexts repeatedly during daily life experiences is helpful for emotion category learning. Prior work has shown that emotion words like “nervous” are heard by young children in naturalistic settings, albeit at relatively low frequencies (Ogren & Sandhofer, 2021). This may explain why explicit emotion labels facilitate learning among 3-year-old children. Namely, since emotion categories are challenging to learn, children *do* likely have some previous experience with these words to draw from. It is possible that having emotion labels paired with vignettes and faces, even in a very brief experiment, was enough to help these young children crystallize their previous experience with these emotions and labels to bring together multiple aspects of emotion into cohesive concepts.

Notably, the present study assessed children’s ability to learn about real emotion categories they will encounter (and likely have previously encountered) in the real world. This allowed the present study to draw clear connections to children’s emotion learning, as these are the same emotion categories that children are tasked with learning in their daily experiences. Additionally, our participants did not demonstrate a strong understanding of these emotion categories at pre-test (average pre-test score was 2.79 out of 8), and thus there was still substantial room for learning about these emotions without ceiling effects. That said, this is an important contrast with previous word learning studies showing that novel labels do influence learning for unfamiliar categories (e.g., Son et al., 2012; Vlach & Sandhofer, 2011). Despite children’s overall low performance at pre-test, in the present study, novel labels did not facilitate children’s learning of emotion categories as they have previously for unfamiliar categories. This may be because emotion categories are abstract and more complex than many learned categories in standard word learning studies, and thus during the learning process children must leverage their prior experience with familiar labels (e.g., Slone & Sandhofer, 2017). To further explore this possibility, it may be fruitful for future research to draw direct comparisons between children’s learning in word learning paradigms and emotion learning paradigms. Further, to

demonstrate the value of specific emotion labels, it may be interesting and informative for future work to examine whether children learn emotions under conditions where emotion labels are provided more generally (i.e., the same label is used for a wider range of faces) as opposed to the specific nature in which they were provided in the present design.

Although these findings may help clarify how emotion labels help children learn about emotions, it is important to note limitations of the present design. First, this study was conducted entirely over Zoom, and children may respond differently to emotional stories and words presented during in-person interactions. However, a recent meta-analysis suggests that developmental studies conducted online are typically comparable to in-person studies (Chuey et al., 2022), and if anything, the effect sizes of online studies may be slightly smaller than those of in-person studies. Therefore, these effects may only be larger if conducted in person. Second, these results only indicate that emotion labels may help connect faces to emotional scenarios for young children. However, emotions are multi-modal and complex (Barrett, 2017), involving factors such as body posture, tone of voice, and knowledge of how people have responded to emotional events in the past. Thus, further research is needed to examine the role of emotion labels in connecting other aspects of emotion beyond just face and scenario. Finally, we did not examine whether the results differed by emotion, as we had limited power to detect such differences. However, it would be useful for future research to examine whether the role of labels on emotion learning differs by specific emotion (for example, for emotions learned earlier versus later in development).

In conclusion, the present study replicates and extends our knowledge of the benefit of emotion labels for early childhood emotion learning. The results replicated the benefit of emotion labels over irrelevant information (Ogren & Sandhofer, 2022) and extended these findings by demonstrating no significant benefit of completely novel labels. Ultimately, these results are beneficial for clarifying the potential role of prior real-world experience with emotion labels on emotion learning.

References

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

Brown, R. (1958). How shall a thing be called? *Psychological Review; 65,* 14–21. doi:10.1037/h0041727

Chuey, A., Boyce, V., Cao, A., & Frank, M. C. (2022). *Conducting developmental research online vs in-person: A meta-analysis.* PsyArXiv doi:10.31234/osf.io/qc6fw

Denham, S. A., Blair, K. A., DeMulder, E., Levitas, J., Sawyer, K., Auerbach-Major, S., & Queenan, P. (2003). Preschool emotional competence: Pathway to social competence? *Child Development, 7,* 238-256. doi:10.1111/1467-8624.00533

Fabes, R. A., Eisenberg, N., Hanish, L. D., & Spinrad, T. L. (2001). Preschoolers' spontaneous emotion vocabulary: Relations to likability. *Early Education and Development*, *12*, 11–27. doi:10.1207/s15566935eed1201_2

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

Fillo, J., Simpson, J. A., Rholes, W. S., & Kohn, J. L. (2015). Dads doing diapers: Individual and relational outcomes associated with the division of childcare across the transition to parenthood. *Journal of Personality and Social Psychology, 108*, 298-316. doi:10.1037/a0038572

Gendron, M., Lindquist, K. A., Barsalou, L., & Barrett, L. F. (2012). Emotion words shape emotion percepts. *Emotion, 12,* 314-325. doi:10.1037/a0026007

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

Gentner, D., & Rattermann, M. J. (1991). Language and the career of similarity. In S. A. Gelman & J. P. Byrnes (Eds.), *Perspectives on thought and language: Interrelations in development* (pp. 225–277). London: Cambridge University Press.

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., Wu, R., LoBue, V., Oakes, L. M., Xu, F., & Barrett, L. F. (2020). Developing an understanding of emotion categories: Lessons from objects. *Trends in Cognitive Sciences, 24,* 39-51. doi:10.1016/j.tics.2019.10.010

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

Hummel, J. E., & Holyoak, K. J. (1997). Distributed representations of structure: A theory of analogical access and mapping. *Psychological Review, 104,* 427–466. doi:10.1037/0033-295X.104.3.427

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

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

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

Ogren, M., & Sandhofer, C. M. (2021). Emotion words in early childhood: A language transcript analysis. *Cognitive Development, 60,* 101122. doi:10.1016/j.cogdev.2021.101122

Ogren, M., & Sandhofer, C. M. (2022). Emotion words link faces to emotional scenarios in early childhood. *Emotion, 22,* 167-178. doi: 10.1037/emo0001063

Ogren, M., & Sandhofer, C. M. (2024, March 11). Emotion Understanding and Labels. Retrieved from osf.io/c3reh

Rattermann, M. J., & Gentner, D. (1998). The effect of language on similarity: The use of relational labels improves young children's performance in a mapping task. In K. Holyoak, D. Gentner, & B. Kokinov (Eds.), *Advances in analogy research: Integration of theory and data from the cognitive, computational, and neural sciences.* Sofia: New Bulgarian University.

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

Slone, L. K. & Sandhofer, C. M. (2017). Consider the category: The effect of spacing depends on individual learning histories. *Journal of Experimental Child Psychology, 159*, 34-49. https://doi.org/10.1016/j.jecp.2017.01.010

Sloutsky, V. M. (2003). The role of similarity in the development of categorization. *Trends in Cognitive Sciences, 7,* 246-251. doi:10.1016/s1364-6613(03)00109-8

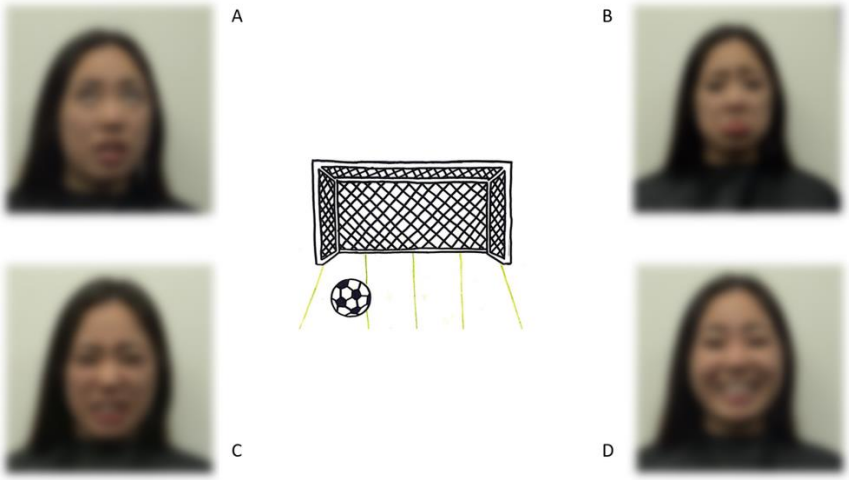
Son, J. Y., Doumas, L. A. A., & Goldstone, R. L. (2010). When do words promote analogical transfer? *The Journal of Problem Solving, 3,* 52-92. doi:10.7771/1932-6246.1079

Son, J. Y., Smith, L. B., Goldstone, R. L., & Leslie, M. (2012). The importance of being interpreted: Grounded words and children's relational reasoning. *Frontiers in Psychology, 3,* Article 45. doi:10.3389/fpsyg.2012.00045

Vlach, H. A., & Sandhofer, C. M. (2011). Developmental differences in children's context-dependent word learning. *Journal of Experimental Child Psychology, 108,* 394-401. doi: 10.1016/j.jecp.2010.09.011

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

Widen, S. C., & Russell, J. A. (2010). Differentiation in preschoolers' categories of emotion. *Emotion, 10,* 651-661. doi: 10.1037/a0019005



"Evelyn is playing soccer. She told her team that she wanted to be the goalie today, but someone else took the spot first. Evelyn stomped away to talk to her coach."

Figure 1. An example of a possible test trial (could be used at pre- or post-test). Text below the image in quotation marks indicates the story accompanying the image and was presented auditorily (i.e., the story text was not presented on the screen to the child). Faces are blurred to prevent recognition of the individual.

"Zoey is going camping with her family. She has never done this before. She will be sleeping in a tent outside and doesn't like that it is very dark."

Figure 2. An example of a possible learning trial. Text below the image in quotation marks indicates the story accompanying this image, and was presented auditorily (i.e., the story text was not presented on the screen to the child). Face is blurred to prevent recognition of the individual.

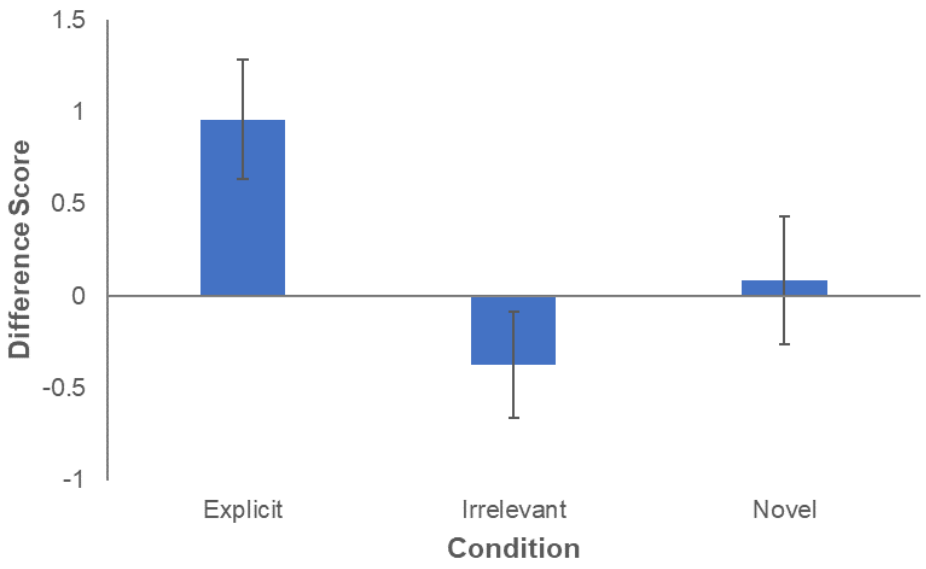


Figure 3. Difference score (post-test minus pre-test) for all three conditions. Error bars represent Standard Error.