



Category Flexibility in Emotion Learning

Rista C. Plate¹ · Kristina Woodard² · Seth D. Pollak²

Received: 8 September 2022 / Accepted: 12 May 2023 / Published online: 9 June 2023
© The Society for Affective Science 2023

Abstract

Learners flexibly update category boundaries to adjust to the range of experiences they encounter. However, little is known about whether the degree of flexibility is consistent across domains. We examined whether categorization of social input, specifically emotions, is afforded more flexibility as compared to other biological input. To address this question, children (6–12 years; 32 female, 37 male; 7 Hispanic or Latino, 62 not Hispanic or Latino; 8 Black or African American, 14 multi-racial, 46 White, 1 selected “other”) categorized faces morphed from calm to upset and animals morphed from a horse to a cow across task phases that differed in the distribution of stimuli presented. Learners flexibly adjusted both emotion and animal category boundaries according to distributional information, yet children showed more flexibility when updating their category boundaries for emotions. These results provide support for the idea that children—who must adjust to the vast and varied emotional signals of their social partners—respond to social signals dynamically in order to make predictions about the internal states and future behaviors of others.

Keywords Emotional development · Emotion categorization · Unsupervised learning · Category flexibility

Children are flexible learners, allowing them to adapt to variable or changing features of their environments. Learning mechanisms that support flexibility, such as attending to underlying statistical relationships, have been examined—and demonstrated—across a variety of domains (Frost et al., 2019). And, there is growing evidence to suggest that such mechanisms also underlie learning in social domains, such as emotional development (Plate et al., 2022; Ruba et al., 2022). For example, children attend to probabilistic patterns in order to make inferences about how different people are feeling (Plate, Wood, et al., 2018; Wu et al., 2017). Furthermore, some domain-general mechanisms appear particularly well suited to aspects of the socio-emotional environment, which is characterized by high variability (Berkay & Jenkins, 2023). One key feature of the variability in children’s social worlds is that emotional input is dynamic. Emotion

signals—i.e., cues that may indicate how someone is feeling, their behavioral intentions, or a response to a previous experience—vary as children seamlessly move between environments (e.g., walk from the classroom to the playground; Aviezer et al., 2008, 2017; Leitzke & Pollak, 2016), interact with different individuals (Friedman et al., 1980; Kring & Gordon, 1998), and cross spaces with culture-specific norms (Cordaro et al., 2018; Niedenthal et al., 2017). Therefore, children must be able to fluidly adapt to the emotional environment in order to make predictions and enact appropriate behaviors (Ruba & Pollak, 2020).

When children encounter emotion signals that differ in expressive intensity, they adjust their criteria for identifying emotion cues based on the underlying statistical distribution (Plate, Wood et al., 2018; Woodard et al., 2021, 2022). Flexibility in the emotion domain mirrors flexibility in other domains, such as color perception (Levari et al., 2018), categorization of novel objects (Kalish et al., 2015), and adjusting to unfamiliar accents (Cristia et al., 2012; Schmale et al., 2012), suggesting that emotion shares features with other informative signals (Ruba et al., 2022; Wu et al., 2020). However, little is understood about the relative degree of flexibility across these instances. Various proposals suggest that the extent to which learners track statistical patterns depends upon the consistency of the input (Frost et al., 2015;

Handling Editor: Eric A Walle

✉ Rista C. Plate
rplate@sas.upenn.edu

¹ Department of Psychology, University of Pennsylvania, 3720 Walnut St, Philadelphia, PA 19104, USA

² Department of Psychology, University of Wisconsin-Madison, 1202 West Johnson Street, Madison, WI 53706, USA

Siegelman et al., 2018). Yet, emotional input is noisy, variable, and subject to referential ambiguity (Ruba et al., 2022). For example, similar situations might result in different emotion experiences or signaling between people or even within the same person across time. Therefore, it is possible that there may be differences in the degree of category flexibility between emotions versus other biologically relevant stimuli.

Current Study and Hypotheses

In the present experiment, participants categorized stimuli that were morphed between facial configurations of emotions (i.e., from calm to upset) and that were morphed between animal configurations (i.e., from a horse to a cow). We employed animals as a contrast biological category because there is developmental change across early to middle childhood in animal categorization. In particular, younger children are less accepting of within-animal variation (Emmons & Kelemen, 2015) and more highly value characteristic properties of animals in categorization (Foster-Hanson & Rhodes, 2019), though children as young as 4 years of age can consider multiple aspects of animals (e.g., prior knowledge, perceptual similarity) in categorization (Carmichael & Hayes, 2001; Hayes, 2006).

First, we asked whether there were differences in the category boundaries between emotion and animal stimuli. Consistent with previous research (Plate, Wood et al., 2018; Woodard et al., 2022), we expected participants to show a low threshold for categorizing faces as upset (i.e., because neutral faces may be interpreted a higher in negativity; Tottenham et al., 2012), but we did not have any a priori predictions for the location of the category boundary between the animal stimuli. Second, we examined whether there are domain-related differences in how participants update these boundaries in the presence, and absence, of explicit feedback. We hypothesized that participants would flexibly update category boundaries in both domains, but left open the possibility that participants would show relatively more flexibility in the emotion domain. Such a finding would be consistent with data indicating high variability in emotion signals (Barrett et al., 2019; Hoemann et al., 2020; Ruba et al., 2022), which may necessitate flexible responding. In addition to domain-related differences, we measured developmental differences to explore whether flexibility varies across middle childhood. Middle childhood is a time when emotion skills, such as recognition (Widen, 2013), differentiation (Nook et al., 2018), and abstraction (Nook et al., 2020) show active development as well as a time when category learning is developing (Hammer et al., 2009; Huang-Pollock et al., 2011). We hypothesized that flexibility would be higher with age, consistent with findings that learning

based on probabilistic information improves with age (Montag, 2021; Plate et al., 2018; Plate, Wood et al. 2018).

Method

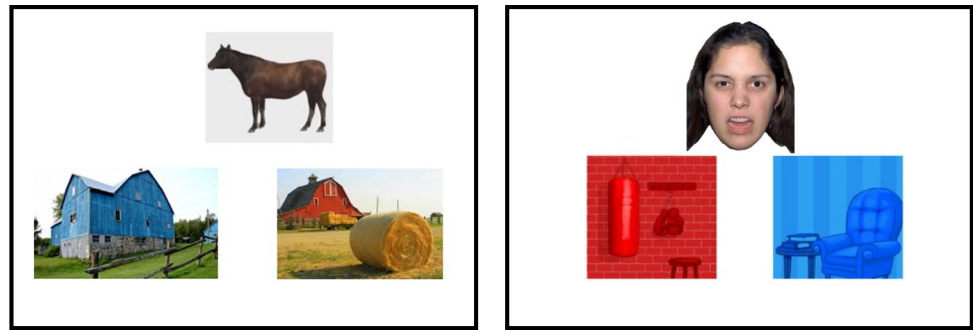
Participants

Participants were 69 children ages 6–12 years (46% female, 54% male, $M_{\text{age}} = 9.82$ years, $SD_{\text{age}} = 1.87$ years). The age range was selected because children as young as 6 years old have previously been able to complete a similar task and reliably categorize ambiguous animal images (Carmichael & Hayes, 2001; Hayes, 2006). The sample size was chosen to be twice that of previous research, which indicated high power (near 100%) for detecting effects of emotion category flexibility, to account for the wider age range used in the present study as compared to previous studies with more limited age ranges (Plate, Wood, et al., 2018; Woodard et al., 2022). One additional participant was excluded for not completing the experiment. Participants were recruited from a mid-size city in the Midwestern region of the USA (10% Hispanic or Latino, 90% not Hispanic or Latino; 12% Black or African American, 20% multiracial, 67% White, 1% selected “other”). Children ages 6–10 years old gave informed oral assent for participation, 11- and 12-year-old children gave written assent, and the parents of all participants provided informed written consent. Children chose a prize and parents received \$20 for study participation. The Institutional Review Board approved the research.

Stimuli and Design

Emotion stimuli included images of facial configurations of one female model (model 10) from the MacArthur Network Face Stimuli Set (Tottenham et al., 2009). The model’s facial configurations were morphed in increments of 5% from a calm face (0% upset) to a 100% upset face (Gao & Maurer, 2009). We included calm and upset morphs to be able to integrate the present research with well-established findings of category flexibility in these stimuli (Plate, Wood et al., 2018; Woodard et al., 2021, 2022). The morphs were labeled as “neutral” and “angry” in the original stimulus set (Tottenham et al., 2009), but we use the terms “calm” and “upset” (in the task and throughout the manuscript) as broad categories and to avoid specific emotion labels. Animal stimuli included images of a horse morphed into a cow. The animal was morphed in increments of 5% from a horse (0% cow) to a 100% cow image and has also been validated in previous research (Wood et al., 2016; Fig. 1). While animal and emotion features are not analogous, we chose animal stimuli as a comparison for a number of reasons. Animals are biological creatures that are subject to variations that

Fig. 1 Example Trial for Each the Animal and Emotion Condition. *Note.* Facial stimuli were obtained, with permission, from Tottenham et al. (2009) and animal stimuli were from Wood et al. (2016). The model depicted here is model #03 from Tottenham et al. (2009); model #10 was used in study (not depicted due to copyright regulations)



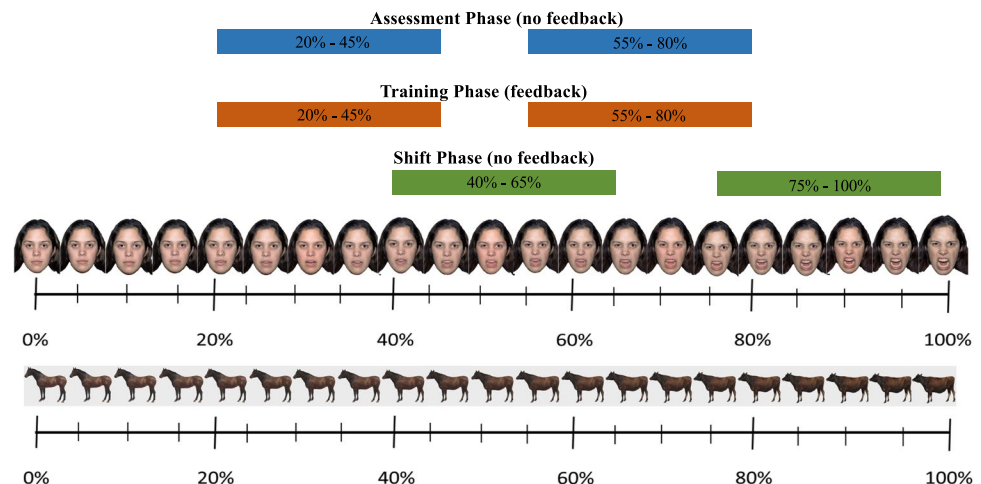
occur within individuals (e.g., an animal's weight may fluctuate over time), across individuals (e.g., individual horses may share some features, such as coloring or shape, but not others), and across geographic contexts (e.g., breeds commons to different areas may vary in features). Participants completed the task twice, once each for the emotion and animal conditions. The order of conditions was randomly assigned, and there was no effect of order on categorizations in any phase ($ps > .50$). The task was adapted from Plate, Wood et al., (2018) and followed the same procedure for the emotion and animal conditions, differing only in the stimulus set used. Participants completed four task phases for each condition (Fig. 2). The practice phase introduced participants to the stimuli and task. The assessment phase was included to measure participants' initial category boundaries. Specifically, we capture the inferences children might make when encountering a non-skewed distribution of emotion/animal signals. The training phase provided a test of whether participants would update their category boundary based on feedback, and was intended to train participants to a centralized category boundary. The training phase captures how children respond when learning whether an inference is accurate. Finally, the shift phase provided an opportunity to observe flexibility of category boundaries in the absence of feedback. Here, we sought to simulate an environment

in which the emotion signals are biased, as children might encounter in learning environments following specific norms or environments that are atypical (e.g., an abusive household in which anger is prevalent (Plate et al., 2019)).

Procedure

Participants sat next to an experimenter and completed the task on the computer. Prior to the start of the task, participants used a slider to view the entire range of morphs for the stimuli for as long as they wanted, with the instructions that they could see the ranges of what Jane's faces look like, or, in the animal condition, the animals on the farm. Next, the experimenter explained how the images could be categorized. In the emotion condition, they said, "When Jane is feeling upset, she likes to go to the red room and practice boxing. When Jane is feeling calm, she likes to go to the blue room and read a book. We need you to help us figure out which room Jane should go to. If you think Jane is feeling upset, click the red button to put her in the red room. If you think Jane is feeling calm, click the blue button to put her in the blue room." In the animal condition, horses and cows were presented with a red barn and blue barn, respectively. The experimenter explained: "Horses on the farm like to go to the red barn to eat hay. Cows like to go to the blue

Fig. 2 Summary of Task Phases. *Note.* Colored boxes indicate morph range sampled during the respective task phase. Facial stimuli were obtained, with permission, from Tottenham et al. (2009) and animal stimuli were from Wood et al. (2016). The model depicted here is model #03 from Tottenham et al. (2009); model #10 was used in study (not depicted due to copyright regulations)



barn to eat grass. We need you to help us figure out which barn to lead each of the animals to. If you think the animal is a horse, click the red button to lead the animal to the red barn. If you think the animal is a cow, click the blue button to lead it to the blue barn.” In six practice trials, participants categorized stimuli that were labeled with their category (i.e., “calm” or “upset” or “horse” or “cow”; 0%, 10%, 20%, 80%, 90%, 100% morphs), received feedback, and repeated trials until correct.

Following the practice trials, participants completed the “assessment phase” in order to measure initial categorization of the stimuli. All participants saw five repetitions of 12 morphs ranging from 20% to 80% upset/cow within each task. Stimuli appeared in a random order in this phase and the following 2 phases; however, each morph was seen once before repetition. The 50% morph was omitted to create a category boundary at the midpoint of the stimuli distribution. Participants did not receive feedback on the accuracy of the categorizations. In this phase and in the following two phases, the words “calm/upset” and “horse/cow” did not appear.

In the subsequent “training phase,” participants saw the same five repetitions of 12 morphs as in the assessment phase (20%–80% upset/cow, 50% morph was omitted). In this phase, participants also received feedback (check mark and “Correct,” or “X” and “Incorrect”) regarding the categorizations. Correct responses were defined as stimuli less than 50% upset/cow placed in the blue room/red barn and stimuli greater than 50% upset/cow placed in the red room/blue barn.

In the final “shift phase,” participants saw five repetitions of 12 morphs in 5% increments so that participants were exposed to the same number of stimuli in the training phase. However, participants saw morphs ranging from 40% to 100% upset/cow in each task. The 70% morph was omitted to create a category boundary at this distribution for consistency (in number and range of morphs) with the earlier phases of the task. Participants did not receive feedback during this phase. This distribution of stimuli has been shown in previous research to engender shifting of category boundaries in facial stimuli in children (Plate, Wood et al., 2018; Woodard et al., 2022).

Analysis

We used logistic mixed effects models and regressed participant selection (i.e., red room for emotion condition and blue barn for the animal condition = 1, blue room/red barn = 0) on the percent morph (mean centered and divided by 10 to facilitate convergence), condition (emotion = -.5, animal = .5), and participant age (mean centered) and all interactions. We included a by-participant random intercept and a random slope for the interaction between percent morph

and condition and lower order effects. If the maximal random effects structure resulted in singular fit, we simplified according to Brauer and Curtin (2018). We provide odds ratios (ORs) and confidence intervals around the ORs as indices of effect size.

Transparency and Openness

The experimental task, analysis script, and de-identified data are available on Open Science Framework: <https://osf.io/zckxq/>. We used R version 3.6.3 for the analyses (R Core Team, 2019), including the tidyverse package (Wickham et al., 2019) for data organization, lme4 package (Bates et al., 2015) for logistic mixed effect models, and ggplot2 (Wickham, 2016) for visualization. The experimental task was programmed in and hosted on PsychoPy (v1.83.04; Peirce et al., 2019). We report all data exclusions, all manipulations, and all measures in the study. The study was not preregistered.

Results

Category Boundaries More Distinguished in Emotion Domain and with Age

First, we examined categorizations in the assessment phase as an initial measure of children’s category boundaries. Consistent with previous research (Plate, Wood, et al., 2018; Woodard et al., 2021, 2022), there was a main effect of percent morph ($b = 2.00$, $z = 21.90$, $p < .001$, $OR = 7.42$, 95% CI [6.20–8.88]; Fig. 3). Participants were more likely to send upset faces to the red room and cows to the blue barn with increasing morph intensity; this is a simple test that children were sensitive to the morph distribution of the stimuli. Next, there was a main effect of condition (animal versus emotion; $b = -2.61$, $z = -10.81$, $p < .001$, $OR = 0.07$, 95% CI [0.05–0.12]). Participants were more likely to select the red room option for upset faces than they were to select the blue barn option for cows: in other words, children had a lower threshold for detecting upset faces. There was also a percent morph-by-condition interaction ($b = -0.74$, $z = -4.75$, $p < .001$, $OR = 0.48$, 95% CI [0.35–0.65]). Participants had less steep category boundaries in the animal versus emotion condition, which indicates that participants were more consistent in their emotion, versus animal, categorizations.

Next, we report on the effect of age. While the main effect of age was not significant ($p = .422$), there were two interactions involving age. First, a percent morph-by-age interaction indicated that category boundaries became steeper with age ($b = 0.16$, $z = 3.72$, $p < .001$, $OR = 1.17$, 95% CI [1.08–1.27]), indicating more consistent categorization choices with age. Second, a condition-by-age interaction

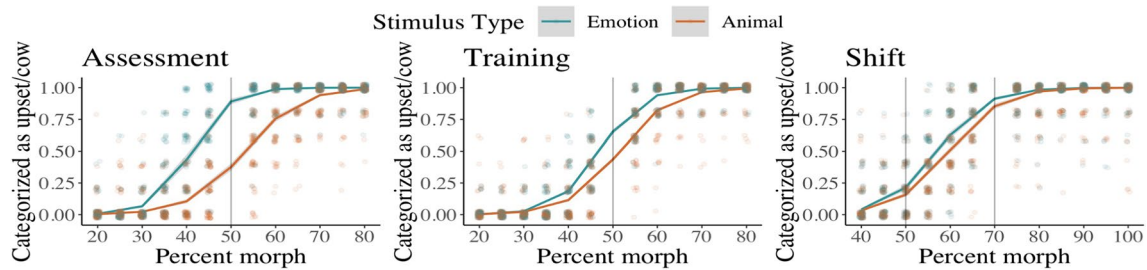


Fig. 3 Participant Responses by Condition for each the Assessment, Training, and Shift Phases. *Note.* Points represent individual-level responses and error bands are standard error. The midpoints of the distributions (50% for the assessment and training phases and 70% for the shift phase) are indicated by vertical lines. The key condition comparisons by phase: (1) in the assessment phase, children

had a lower threshold for categorizations in the emotion versus animal condition, but steeper boundaries in the emotion condition. (2) In the training phase, children's categorizations moved closer to the midpoint in the emotion condition. (3) In the shift phase, children changed their category boundaries to a greater extent in the emotion condition

indicated that children identified the model as upset at a lower threshold with age ($b = -0.23$, $z = -1.94$, $p = .05$, $OR = 0.80$, 95% CI [0.63–1.00]). The three-way interaction between percent morph, condition, and age was not significant ($p = .81$). We additionally checked whether our other demographic variables (race and gender) related to category boundaries. We find that race was not related to categorization ($X^2(5) = 2.92$, $p = .71$). However, girls had a slightly lower threshold for categorizing faces as upset as compared with boys ($b = -.12$, $X^2(1) = 3.77$, $p = .05$). Together, these results show consistent categorization patterns in each the animal and emotion conditions, but highlight that category boundary thresholds and steepness differ by domain and across development.

Category Boundaries Sensitive to Feedback Across Domains and Age

When participants received feedback on their categorization choices, the patterns of findings mirrored the results of the assessment phase (Table 1). In other words, explicit feedback regarding how the stimuli should be categorized did not eliminate condition or age differences. Because these differences persisted in the training phase, we included phase as a predictor in the model (assessment = $-.5$, training = $.5$) to examine whether there were differences in the extent

to which participants changed their category boundaries between the assessment and training phases based on condition and age. There was a main effect of phase ($b = -0.36$, $z = -5.72$, $p < .001$, $OR = 0.70$, 95% CI [0.62–0.79]), indicating that feedback was effective for shifting category boundaries toward the midpoint of the distribution. A percent morph-by-phase interaction ($b = 0.32$, $z = 6.12$, $p < .001$, $OR = 1.38$, 95% CI [1.24–1.53]) indicated category boundaries becoming steeper during the training phase, again suggesting that children used the feedback to inform their categorization choices. Finally, a condition-by-phase interaction showed that there is more movement toward the midpoint in the emotion condition relative to the animal condition ($b = 1.08$, $z = 8.64$, $p < .001$, $OR = 2.95$, 95% CI [2.31–3.77]). The interaction between phase and age, and higher order interactions including phase were not significant ($ps > .11$). Together these results indicate that children more readily updated emotion category boundaries (as compared to animal category boundaries) in response to feedback.

Category Boundaries More Flexible in Emotion Domain and Across Age

Because the difference between the emotion and animal conditions remained after the training phase, we again included

Table 1 Full Model Output for the Training Phase

Predictors	b	z	p	OR	95% CI
Percent morph	1.95	40.92	< .001	7.04	6.41–7.73
Condition	– 0.89	– 8.67	< .001	0.41	0.34–0.50
Age	0	– 0.1	.917	1	0.94–1.06
Percent morph-by-condition	– 0.33	– 2.56	.01	0.72	0.56–0.92
Percent morph-by-age	0.18	8.86	< .001	1.19	1.15–1.24
Condition-by-age	– 0.09	– 1.83	.067	0.91	0.83–1.01
Percent morph-by-condition-by-age	0.01	0.12	.907	1.01	0.89–1.15

phase (training = $-.5$, shift = $.5$) in the model to best capture how categorizations changed when participants went from the training phase to categorizing a shifted selection of morphs in the shift phase. First, there were a number of main effects, including main effect of percent morph ($b = 1.84$, $z = 24.01$, $p < .001$, $OR = 6.32$, 95% CI [5.44–7.35]) and condition ($b = -.059$, $z = -6.04$, $p < .001$, $OR = 0.55$, 95% CI [0.46–0.67]), consistent with the preceding phases. There was also a main effect of phase ($b = -1.53$, $z = -14.89$, $p < .001$, $OR = 0.22$, 95% CI [0.18–0.27]) such that participants were shifting toward the center of the new distribution and a main effect of age ($b = -.07$, $z = -2.55$, $p = .01$, $OR = 0.93$, 95% CI [0.88–0.98]) with older children more likely to categorize upset/cow morphs as going to the red room/blue barn.

The effects of interest involving condition and participant age can be further understood in the context of interactions. Consistent with the prediction that the emotion domain necessitates greater flexibility, participants changed their category boundaries between the training and shift phases to a greater extent in the emotion versus farm condition ($b = 0.49$, $z = 3.85$, $p < .001$, $OR = 1.63$, 95% CI [1.27–2.09]; Fig. 3). As in the preceding phases, category boundaries were also steeper in the emotion condition ($b = -.027$, $z = -5.60$, $p < .001$, $OR = 0.76$, 95% CI [0.69–0.84]). However, the difference in steepness was somewhat smaller in the shift phase ($b = 0.19$, $z = 1.94$, $p = .05$, $OR = 1.21$, 95% CI [1.00–1.46]), and category boundaries in both conditions were less steep in the shift versus training phase ($b = -.047$, $z = -9.64$, $p < .001$, $OR = 0.63$, 95% CI [0.57–0.69]).

Finally, regarding developmental differences, there was an age-by-phase interaction ($b = -.013$, $z = -2.52$, $p = .01$, $OR = 0.88$, 95% CI [0.79–0.97]), suggesting that participants updated their category boundaries more between phases with age; similar to the previous phases, the steepness of the category boundaries also increased with age ($b = 0.18$, $z = 4.45$, $p < .001$, $OR = 1.19$, 95% CI [1.10–1.29]). All other effects in the model were not significant ($ps > .21$).

Discussion

Our primary aim was to examine potential domain-relevant differences in the flexibility of categorization. Children adjusted their categorizations in response to the statistical distribution of emotion-relevant signals to a greater extent than other biological stimuli (i.e., animals). These findings extend our knowledge of human learning in two ways. First, the data add to a growing body of research providing evidence for domain-general learning mechanisms underlying children's mastery of the vast and varied social input that comprise emotional development (Plate et al., 2022; Ruba et al., 2022). Second, the results from this study highlight

how quickly learners adapt to variability in the emotional domain (here within 60 trials), leveraging statistical regularities in the environment to make predictions about others. We also observed steeper category boundaries for emotion, relative to animal, stimuli; this indicates more consistency and adept flexibility in children's responses to emotion-related stimuli.

Emotion categories were more flexible both in the presence and absence of feedback, even without the use of emotion labels. Perceivers do encounter labeled emotions, for example, in books or when someone verbally expresses their emotion state (Ahn, 2005; Pollak & Thoits, 1989). However, there is variation in how much these labels facilitate emotion categorization (Hoemann et al., 2022) and much emotional input is unlabeled (see Ruba et al., 2022 for a number of examples). Feedback and statistical structure may meaningfully interact to support children's emotional development in addition to, or in conjunction with, labeling (Plate et al., 2022). More research is needed to understand when feedback is necessary, or sufficient, for emotion knowledge acquisition. Our instructions provided the labels "upset" and "calm" to reference mental states. We selected these words to (a) avoid specific emotion categories (such as angry or happy), (b) encourage children to engage holistically with the stimuli rather than cue specific features (such as a scowl), or (c) concretely attend to one variant or exemplar of the faces. Mental state content includes some level of uncertainty and dynamicity (Berkay & Jenkins, 2023), and it is possible that these features cued flexibility.

Older children showed more category flexibility, regardless of the domain. On the one hand, it is not surprising that older children demonstrated behavior consistent with task performance given that both emotion recognition (Widen, 2013; Widen & Russell, 2008) and category learning (Hammer et al., 2009; Huang-Pollock et al., 2011) develop substantially during this age period. On the other hand, research suggests that younger children tend to show more flexibility in learning contexts (Lucas et al., 2014). The present results are consistent with the idea that flexibility in response to social cues may be observed later in development as children spend more time in peer contexts (Gopnik et al., 2017).

The findings should be considered alongside several limitations that provide guidance for future research. First, emotion signals were compared to one other visual category (animals), and animals and emotions are not equivalent on a number of dimensions. For example, while there is variability both within and across animal forms (as with emotion configurations), there is less dynamic change between species than emotion signals. And, while perceivers make emotion judgments categorically (Etcoff & Magee, 1992), as they are likely to do with animals as well, it is possible that perceivers are more attuned to the continuous features of emotion signals. For example, it could be more important

to know if someone is enraged versus mildly irritated than to know if an animal more closely resembles a cow versus a horse. It is possible that flexibility in response to emotion signals might be more similar in other domains such as language (Ruba et al., 2022; Weatherholtz & Jaeger, 2016) with which emotion closely shares features of variability. For instance, children use statistical information to adjust to a variety of vocal cues including nonverbal emotion vocalizations (i.e., a yell) and emotionally inflected speech (Woodard et al., 2021). Nevertheless, we included animals as stimuli because they do share many characteristics with emotion signals.

Perhaps counterintuitively, when participants encountered more anger signals, they raised their threshold for categorizing faces as angry. This finding raises considerations for contexts in which participants might be exposed to biased emotion signals. For example, there is evidence to suggest that exposure to environments with heightened anger cues (Plate et al., 2019) is associated with lower thresholds for anger detection (Pollak & Kistler, 2002). Therefore, multiple mechanisms likely underly shifting in complex social environments. Future research could also explore emotion signals beyond anger—which may be particularly salient given its role in threat detection—and expand to facial signals that are not readily considered “basic emotions.” Finally, the stimuli were impoverished in comparison to the complex environments in which perceivers typically view emotions (Barrett et al., 2019). Future research should consider how additional cues in the environment affect category flexibility across domains.

This experiment highlights domain differences in category flexibility and points to emotion as a domain in which children evidence particular flexibility in adjusting to environmental input. This makes sense in that variability is a key feature of emotions, with meaningful differences across and within individuals and contexts. Being able to quickly and accurately adapt to this variability promotes social competence by allowing children to make reasonable inferences about what others may be feeling, and to predict the behavior of other people based on those inferences.

Acknowledgements We thank the families who participated in this study and the research assistants who helped conduct the research, particularly Cassandra Windau, Carolyn Meissner, and Emily Weiss. Images were reproduced with permission from Tottenham et al. (2009) and Wood et al. (2016). The experiment was approved by the Institutional Review Board.

Additional Information

Funding This research was supported by the National Institute of Mental Health through grant R01MH61285 to S.D.P. and in part by a core grant to the Waisman Center from the National Institute of Child Health and Human Development (P50HD105353). R.C.P. was

supported by a National Science Foundation Graduate Research Fellowship (DGE-1256259) and the Richard L. and Jeanette A. Hoffman Wisconsin Distinguished Graduate Fellowship.

Conflict of Interest The authors declare competing interest.

Data Availability The experimental paradigm, de-identified data, and analysis scripts are available on Open Science Framework: <https://osf.io/zckxq/>.

Author Contribution S.D.P. acquired funding for the research. R.C.P. and K.W. conducted data curation and formal analysis. R.C.P. prepared the original draft of the manuscript. All authors were involved with conceptualization, methodology, and reviewing and editing.

Ethical Approval The Institutional Review Board approved the research.

Informed Consent The parents of all participants provided informed written consent.

References

- Ahn, H. J. (2005). Teachers? Discussions of emotion in child care centers. *Early Childhood Education Journal*, 32(4), 237–242. <https://doi.org/10.1007/s10643-004-1424-6>
- Aviezer, H., Hassin, R. R., Ryan, J., Grady, C., Susskind, J., Anderson, A., Moscovitch, M., & Bentin, S. (2008). Angry, disgusted, or afraid? Studies on the malleability of emotion perception. *Psychological Science*, 19(7), 724–732. <https://doi.org/10.1111/j.1467-9280.2008.02148.x>
- Aviezer, H., Ensenberg, N., & Hassin, R. R. (2017). The inherently contextualized nature of facial emotion perception. *Current Opinion in Psychology*, 17, 47–54. <https://doi.org/10.1016/j.copsyc.2017.06.006>
- 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), 1–68. <https://doi.org/10.1177/1529100619832930>
- Bates, D., Mächler, M., Bolker, B., & Walker, S. (2015). Fitting linear mixed-effects models using lme4. *Journal of Statistical Software*, 67(1). <https://doi.org/10.18637/jss.v067.i01>
- Berkay, D., & Jenkins, A. C. (2023). A role for uncertainty in the neural distinction between social and nonsocial thought. *Perspectives on Psychological Science*, 18(2), 491–502. <https://doi.org/10.1177/17456916221112077>
- Brauer, M., & Curtin, J. J. (2018). Linear mixed-effects models and the analysis of nonindependent data: A unified framework to analyze categorical and continuous independent variables that vary within-subjects and/or within-items. *Psychological Methods*, 23(3), 389–411. <https://doi.org/10.1037/met0000159>
- Carmichael, C. A., & Hayes, B. K. (2001). Prior knowledge and exemplar encoding in children’s concept acquisition. *Child Development*, 72(4), 1071–1090. <https://doi.org/10.1111/1467-8624.00335>
- Cordaro, D. T., Sun, R., Keltner, D., Kamble, S., Huddar, N., & McNeil, G. (2018). Universals and cultural variations in 22 emotional expressions across five cultures. *Emotion*, 18(1), 75–93. <https://doi.org/10.1037/emo0000302>
- Cristia, A., Seidl, A., Vaughn, C., Schmale, R., Bradlow, A., & Flocia, C. (2012). Linguistic processing of accented speech across the lifespan. *Frontiers in Psychology*, 3, 479. <https://doi.org/10.3389/fpsyg.2012.00479>

- Emmons, N. A., & Kelemen, D. A. (2015). Young children's acceptance of within-species variation: Implications for essentialism and teaching evolution. *Journal of Experimental Child Psychology*, 139, 148–160. <https://doi.org/10.1016/j.jecp.2015.05.011>
- Etcoff, N. L., & Magee, J. J. (1992). Categorical perception of facial expressions. *Cognition*, 44(3), 227–240. [https://doi.org/10.1016/0010-0277\(92\)90002-Y](https://doi.org/10.1016/0010-0277(92)90002-Y)
- Foster-Hanson, E., & Rhodes, M. (2019). Is the most representative skunk the average or the stinkiest? Developmental changes in representations of biological categories. *Cognitive Psychology*, 110, 1–15. <https://doi.org/10.1016/j.cogpsych.2018.12.004>
- Friedman, H. S., Prince, L. M., Riggio, R. E., & DiMatteo, M. R. (1980). Understanding and assessing nonverbal expressiveness: The Affective Communication Test. *Journal of Personality and Social Psychology*, 39(2), 333–351. <https://doi.org/10.1037/0022-3514.39.2.333>
- Frost, R., Armstrong, B. C., Siegelman, N., & Christiansen, M. H. (2015). Domain generality versus modality specificity: The paradox of statistical learning. *Trends in Cognitive Sciences*, 19(3), 117–125. <https://doi.org/10.1016/j.tics.2014.12.010>
- Frost, R., Armstrong, B. C., & Christiansen, M. H. (2019). Statistical learning research: A critical review and possible new directions. *Psychological Bulletin*, 145(12), 1128–1153. <https://doi.org/10.1037/bul0000210>
- Gao, X., & Maurer, D. (2009). Influence of intensity on children's sensitivity to happy, sad, and fearful facial expressions. *Journal of Experimental Child Psychology*, 102(4), 503–521. <https://doi.org/10.1016/j.jecp.2008.11.002>
- Gopnik, A., O'Grady, S., Lucas, C. G., Griffiths, T. L., Wente, A., Bridgers, S., Aboody, R., Fung, H., & Dahl, R. E. (2017). Changes in cognitive flexibility and hypothesis search across human life history from childhood to adolescence to adulthood. *Proceedings of the National Academy of Sciences*, 114(30), 7892–7899. <https://doi.org/10.1073/pnas.1700811114>
- Hammer, R., Diesendruck, G., Weinshall, D., & Hochstein, S. (2009). The development of category learning strategies: What makes the difference? *Cognition*, 112(1), 105–119. <https://doi.org/10.1016/j.cognition.2009.03.012>
- Hayes, B. K. (2006). Knowledge, development, and category learning. In *Psychology of learning and motivation* (Vol. 46, pp. 37–77). Academic Press. [https://doi.org/10.1016/S0079-7421\(06\)46002-3](https://doi.org/10.1016/S0079-7421(06)46002-3)
- 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(1), 39–51. <https://doi.org/10.1016/j.tics.2019.10.010>
- Hoemann, K., Gendron, M., & Barrett, L. F. (2022). Assessing the power of words to facilitate emotion category learning. *Affective Science*, 3(1), 69–80. <https://doi.org/10.1007/s42761-021-00084-4>
- Huang-Pollock, C. L., Maddox, W. T., & Karalunas, S. L. (2011). Development of implicit and explicit category learning. *Journal of Experimental Child Psychology*, 109(3), 321–335. <https://doi.org/10.1016/j.jecp.2011.02.002>
- Kalish, C. W., Zhu, X., & Rogers, T. T. (2015). Drift in children's categories: When experienced distributions conflict with prior learning. *Developmental Science*, 18(6), 940–956. <https://doi.org/10.1111/desc.12280>
- Kring, A. M., & Gordon, A. H. (1998). Sex differences in emotion: expression, experience, and physiology. *Journal of Personality and Social Psychology*, 74(3), 686–703. <https://doi.org/10.1037/0022-3514.74.3.686>
- Leitzke, B. T., & Pollak, S. D. (2016). Developmental changes in the primacy of facial cues for emotion recognition. *Developmental Psychology*, 52(4), 572–581. <https://doi.org/10.1037/a0040067>
- Levari, D. E., Gilbert, D. T., Wilson, T. D., Sievers, B., Amodio, D. M., & Wheatley, T. (2018). Prevalence-induced concept change in human judgment. *Science*, 360(6396), 1465–1467. <https://doi.org/10.1126/science.aap8731>
- Lucas, C. G., Bridgers, S., Griffiths, T. L., & Gopnik, A. (2014). When children are better (or at least more open-minded) learners than adults: Developmental differences in learning the forms of causal relationships. *Cognition*, 131(2), 284–299. <https://doi.org/10.1016/j.cognition.2013.12.010>
- Montag, J. L. (2021). Limited evidence for probability matching as a strategy in probability learning tasks. In *Psychology of learning and motivation* (p. S0079742121000050). Elsevier. <https://doi.org/10.1016/bs.plm.2021.02.005>
- Niedenthal, P. M., Rychlowska, M., & Wood, A. (2017). Feelings and contexts: Socioecological influences on the nonverbal expression of emotion. *Current Opinion in Psychology*, 17, 170–175. <https://doi.org/10.1016/j.copsyc.2017.07.025>
- Nook, E. C., Sasse, S. F., Lambert, H. K., McLaughlin, K. A., & Somerville, L. H. (2018). The nonlinear development of emotion differentiation: Granular emotional experience is low in adolescence. *Psychological Science*, 29(8), 1346–1357. <https://doi.org/10.1177/0956797618773357>
- Nook, E. C., Stavish, C. M., Sasse, S. F., Lambert, H. K., Mair, P., McLaughlin, K. A., & Somerville, L. H. (2020). Charting the development of emotion comprehension and abstraction from childhood to adulthood using observer-rated and linguistic measures. *Emotion*, 20(5), 773–792. <https://doi.org/10.1037/emo0000609>
- Peirce, J., Gray, J. R., Simpson, S., MacAskill, M., Höchenberger, R., Sogo, H., ... & Lindeløv, J. K. (2019). PsychoPy2: Experiments in behavior made easy. *Behavior Research Methods*, 51, 195–203. <https://doi.org/10.3758/s13428-018-01193-y>
- Plate, R. C., Fulvio, J. M., Shutts, K., Green, C. S., & Pollak, S. D. (2018). Probability learning: Changes in behavior across time and development. *Child Development*, 89(1), 205–218. <https://doi.org/10.1111/cdev.12718>
- Plate, R. C., Wood, A., Woodard, K., & Pollak, S. D. (2018). Probabilistic learning of emotion categories. *Journal of Experimental Psychology: General*, 148(10), 1814–1827. <https://doi.org/10.1037/xge0000529>
- Plate, R. C., Bloomberg, Z., Bolt, D. M., Bechner, A. M., Roeber, B. J., & Pollak, S. D. (2019). Abused children experience high anger exposure. *Frontiers in Psychology*, 10, 440. <https://doi.org/10.3389/fpsyg.2019.00440>
- Plate, R. C., Woodard, K., & Pollak, S. D. (2022). Statistical learning in an emotional world. In D. Dukes, A. D. Samson, & E. A. Walle (Eds.), *The Oxford handbook of emotional development* (pp. 78–92).
- Pollak, S. D., & Kistler, D. J. (2002). Early experience is associated with the development of categorical representations for facial expressions of emotion. *Proceedings of the National Academy of Sciences*, 99(13), 9072–9076.
- Pollak, L. H., & Thoits, P. A. (1989). Processes in emotional socialization. *Social Psychology Quarterly*, 52(1), 22. <https://doi.org/10.2307/2786901>
- R Core Team. (2019). *R: A language and environment for statistical computing*. Vienna, Austria: R Foundation for Statistical Computing. <https://www.R-project.org/>
- Ruba, A., & Pollak, S. (2020). The development of emotion reasoning in infancy and early childhood. *Annual Review of Developmental Psychology*, 2, 503–531. <https://doi.org/10.1146/annurev-devpsych-060320-102556>
- Ruba, A. L., Pollak, S., & Saffran, J. (2022). Acquiring complex communicative systems: Statistical learning of language and emotion. *Topics in Cognitive Science*, 14(3), 432–450.
- Schmale, R., Cristia, A., & Seidl, A. (2012). Toddlers recognize words in an unfamiliar accent after brief exposure: Brief exposure to an unfamiliar accent. *Developmental Science*, 15(6), 732–738. <https://doi.org/10.1111/j.1467-7687.2012.01175.x>

- Siegelman, N., Bogaerts, L., Elazar, A., Arciuli, J., & Frost, R. (2018). Linguistic entrenchment: Prior knowledge impacts statistical learning performance. *Cognition*, 177, 198–213. <https://doi.org/10.1016/j.cognition.2018.04.011>
- Tottenham, N., Tanaka, J. W., Leon, A. C., McCarry, T., Nurse, M., Hare, T. A., Marcus, D. J., Westerlund, A., Casey, B., & Nelson, C. (2009). The NimStim set of facial expressions: Judgments from untrained research participants. *Psychiatry Research*, 168(3), 242–249. <https://doi.org/10.1016/j.psychres.2008.05.006>
- Tottenham, N., Phuong, J., Flannery, J., Gabard-Durnam, L., & Goff, B. (2012). A negativity bias for ambiguous facial-expression valence during childhood: Converging evidence from behavior and facial corrugator muscle responses. *Emotion*, 13(1), 92. <https://doi.org/10.1037/a0029431>
- Weatherholtz, K., & Jaeger, T. F. (2016). *Speech Perception and generalization across talkers and accents (Vol. 1)*. Oxford University Press. <https://doi.org/10.1093/acrefore/9780199384655.013.95>
- Wickham, H. (2016). *Ggplot2*. Springer International Publishing. <https://doi.org/10.1007/978-3-319-24277-4>
- Wickham, H., Averick, M., Bryan, J., Chang, W., McGowan, L., François, R., Grolemund, G., Hayes, A., Henry, L., Hester, J., Kuhn, M., Pedersen, T., Miller, E., Bache, S., Müller, K., Ooms, J., Robinson, D., Seidel, D., Spinu, V., et al. (2019). Welcome to the Tidyverse. *Journal of Open Source Software*, 4(43), 1686. 10.21105/joss.01686.
- 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.
- Widen, S. C., & Russell, J. A. (2008). Children acquire emotion categories gradually. *Cognitive Development*, 23(2), 291–312. <https://doi.org/10.1016/j.cogdev.2008.01.002>
- Wood, A., Lupyan, G., Sherrin, S., & Niedenthal, P. (2016). Altering sensorimotor feedback disrupts visual discrimination of facial expressions. *Psychonomic Bulletin & Review*, 23(4), 1150–1156. <https://doi.org/10.3758/s13423-015-0974-5>
- Woodard, K., Plate, R. C., Morningstar, M., Wood, A., & Pollak, S. D. (2021). Categorization of vocal emotion cues depends on distributions of input. *Affective Science*, 2(3), 301–310. <https://doi.org/10.1007/s42761-021-00038-w>
- Woodard, K., Plate, R. C., & Pollak, S. D. (2022). Children track probabilistic distributions of facial cues across individuals. *Journal of Experimental Psychology: General*, 151(2), 506–511. <https://doi.org/10.1037/xge0001087>
- 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(45), 11896–11901. <https://doi.org/10.1073/pnas.1707715114>
- Wu, Y., Schulz, L., Frank, M. C., & Gweon, H. (2020). *Emotion as information in early social learning*. PsyArXiv. <https://doi.org/10.31234/osf.io/fz857>

Springer Nature or its licensor (e.g. a society or other partner) holds exclusive rights to this article under a publishing agreement with the author(s) or other rightsholder(s); author self-archiving of the accepted manuscript version of this article is solely governed by the terms of such publishing agreement and applicable law.