



Examining Preverbal Infants' Ability to Map Labels to Facial Configurations

Ashley L. Ruba¹ · Lasana T. Harris² · Makeba Parramore Wilbourn³

Received: 15 April 2020 / Accepted: 18 August 2020 / Published online: 17 September 2020
© The Society for Affective Science 2020

Abstract

Language is important for emotion perception, but very little is known about *how* emotion labels are learned. The current studies examine how preverbal infants map novel labels onto facial configurations. Across studies, infants were tested with a modified habituation paradigm (“switch design”). Experiments 1 and 2 found that 18-month-olds, but not 14-month-olds, mapped novel labels (“blicket” and “toma”) to *human* facial configurations associated with happiness and sadness. Subsequent analyses revealed that vocabulary size positively correlated with 14-month-olds’ ability to form the mappings. Experiment 3 found that 14-month-olds were able to map novel labels to facial configurations when the visual complexity of the stimuli was reduced (i.e., by using *cartoon* facial configurations). This suggests that cognitive maturation and language development influence infants’ associative word learning with facial configurations. The current studies are a critical first step in determining how infants navigate the complex process of learning emotion labels.

Keywords Language · Infancy · Emotions · Word learning

Language is important for emotion perception in early childhood (Ruba, Meltzoff, & Repacholi, 2020b). The ability to label facial configurations is both an index of emotion understanding (Pons, Harris, & de Rosnay, 2004) and a predictor of social and academic competence (Izard et al., 2001). Consequently, much research has focused on *when* children learn different emotion labels. This is a gradual process that begins late in the second year of life and continues throughout early childhood (Widen, 2013). However, relatively little is known about *how* emotion labels are learned (see Hoemann, Xu, & Barrett, 2019; Ruba & Repacholi, 2019; Shablack, Becker, & Lindquist, 2019). The current studies examine an

important component of emotion label learning: infants’ ability to form associative links between labels and facial configurations.

Associative word learning is foundational for vocabulary development. Infants must first map a label to a referent (e.g., “ball”) before they can attribute symbolic meaning to that label (Werker, Cohen, Lloyd, Casasola, & Stager, 1998). Researchers commonly measure associative word learning with a habituation paradigm called the “switch design” (for a meta-analysis, see Tsui, Byers-Heinlein, & Fennell, 2019), in which infants are habituated to two referent-label pairings (i.e., *object #1*-“blicket”, *object #2*-“toma”) and tested with (a) a *same* trial that maintains one of the pairings (e.g., *object #1*-“blicket”) and (b) a *switch* trial that violates one of the pairings (e.g., *object #1*-“toma”). Infants provide evidence of forming associative links between referents and labels if they look longer to the *switch* trial compared with the *same* trial. Results with this paradigm show that 14-month-olds can map novel labels to objects (e.g., dog, ball) (Fennell & Werker, 2003; Werker et al., 1998) and spatial prepositions (e.g., in, on) (Casasola & Wilbourn, 2004). However, as relatively novice word learners, 14-month-olds have fewer cognitive resources to recruit when faced with more complex associative word learning tasks (Stager & Werker, 1997). For example, 14-month-olds fail to map phonetically similar labels (e.g.,

Handling editor: Kristen Lindquist

Electronic supplementary material The online version of this article (<https://doi.org/10.1007/s42761-020-00015-9>) contains supplementary material, which is available to authorized users.

✉ Ashley L. Ruba
ruba@wisc.edu

¹ Department of Psychology, University of Wisconsin – Madison, Waisman Center 399, 1500 Highland Avenue, Madison, WI, USA

² University College London, London, UK

³ Duke University, Durham, NC, USA

“bih,” “dih”) to objects (Stager & Werker, 1997), only succeeding at 17 months (Werker, Fennell, Corcoran, & Stager, 2002). Likewise, 18-month-olds, but not 14-month-olds, can map novel labels to more complex referents, such as causal actions (e.g., pushing, pulling) (Casasola & Cohen, 2000; Chan et al., 2011).

An open question is whether infants can also form associative links between novel labels and facial configurations. Prior work suggests that 3-year-olds, but not younger 2-year-olds, can use a process-of-elimination strategy to associate a novel label (e.g., “pax”) with a non-emotion facial configuration in a behavioral paradigm (Nelson & Russell, 2016). However, given that 14-month-olds can map novel labels onto objects (Werker et al., 1998) and spatial prepositions (Casasola & Wilbourn, 2004) in the looking time “switch” paradigm, we hypothesized that 14-month-olds would also map novel labels (e.g., “toma,” “blicket”) to facial configurations associated with happiness and sadness.

Experiment 1

Method

Participants The study was performed to ethical standards as laid down in the 1964 Declaration of Helsinki and conducted with approval of the institutional review board at Duke University (Approval Number: B0181, Protocol Title: “Early Language Learning: Labels for Emotional Expressions”). Participants were recruited in the southeastern United States through public birth records and announcements. A power analysis indicated that a sample size of 28 infants would be sufficient to detect reliable differences in a 3-way within-subjects ANOVA, assuming a medium effect size ($\eta_p^2 = .06$) (Tsui et al., 2019) at the .05 alpha level with a power of .80. This was preselected as the stopping rule for the study.

A total of 47 14-month-olds were recruited and tested. The final sample included 29 14-month-old infants (14 female, $M = 13.90$ months, $SD = .30$ months, range = 13.52 months – 14.51 months). Parents identified their infants as White (90%, $n = 26$), Black (7%, $n = 2$), or Asian (3%, $n = 1$). Eighteen additional infants participated in the study but were excluded from final analyses for computer error ($n = 2$), failure to meet the habituation criteria, described below ($n = 5$), extreme looking times at the *Same* test trial (longer than 22.78 s; $n = 3$), suggesting a failure to habituate (Oakes, 2010), inability to see eyes during the experiment ($n = 1$), fussiness or inattentiveness that lead to difficulties with accurate coding ($n = 5$), or parental interference ($n = 2$). Exclusions made for fussiness and inattentiveness were initially made by the blind, online coder, who marked the tested infant as likely too fussy and/or inattentive for coding to be reliable. A second blind coder confirmed this decision during secondary offline reliability

coding. Infants were also excluded from final analyses if their looking time to the *Same* test trial was greater than the sample’s average looking time to the longest four habituation trials (i.e., 22.78 s), suggesting that these infants did not truly habituate (Oakes, 2010). This attrition rate (38%) is similar to previous studies that have used the switch design (Casasola & Cohen, 2000; Casasola & Wilbourn, 2004; Werker et al., 1998).

Although the switch design has a high attrition rate, there are several advantages in using this paradigm (Werker et al., 1998). First, the switch design controls for potential confounds that may influence infant associative word learning, such as environmental cues and adult assistance. Second, by using looking time as the dependent variable, younger, preverbal infants can be tested. For these reasons, the switch design is one of the most common associative word learning paradigms in infancy research and has been used in nearly 150 studies (Tsui et al., 2019).

Stimuli Semi-dynamic events were created in iMovie with visual and auditory stimuli (Fig. 1). Events were semi-dynamic, given that moving stimuli facilitate associative word learning for 14-month-olds (Werker et al., 1998). Visual stimuli were pictures of facial configurations from the Radboud Faces Database (Langner et al., 2010). Each video began with a female child displaying a neutral facial configuration. After 1.5 s, the target facial configuration (e.g., happy) appeared. The target remained for 3.5 s before shifting to a black screen, which lasted for 1 s. Each 6 s progression was looped continuously five times, creating a 30 s trial. To promote success at the task, maximally dissimilar facial configurations (i.e., those associated with happiness and sadness) were selected for the habituation events (see also, Werker et al., 1998). Facial configurations associated with happiness and sadness differ based on both affective dimensions and facial muscle movements and are readily discriminable by infants as young as 5 months of age (see Ruba & Repacholi, 2019).

The auditory stimuli were pre-recorded nonsense words spoken by a native English-speaking female in infant-directed speech. The novel labels (i.e., “blicket,” “toma,” “dax”) are phonemically “legal” in English and have been used in previous infancy studies (Ruba et al., 2020b). In each event, the novel labels were spoken twice after the target facial configuration appeared. The novel label was never presented immediately before or during the appearance of the target. This presentation format increased the likelihood that infants would map the novel label to the target facial configuration and decreased the likelihood that infants would (a) associate the label with the change in facial configuration or (b) make causal attributions (e.g., the label *caused* the facial configuration to change).

Apparatus Infants were tested in a 3 m × 3 m room on their parent’s lap. Infants sat approximately 127 cm from a 50-cm


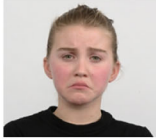








| Habituation Events | | Test Events | | |
|---|---|---|--|---|
| | | Same | Switch | Post |
|  |  |  |  |  |
| “Blicket” | “Toma” | “Toma” | “Toma” | “Dax” |
|  |  |  |  |  |
| “Blicket” | “Toma” | “Toma” | “Toma” | “Dax” |

Fig. 1 Experiment design and example stimuli. Human faces (top row) were used for Experiment 1 and 2. Cartoon faces (bottom row) were used for Experiment 3. Habituation pairings were counterbalanced across participants. The presentation order of the *Same* and *Switch* trials was also counterbalanced. Habituation and test events were presented in a randomized order, except for the *Post* trial, which was always presented

last. Pictures reprinted with permission from the creators of the Radboud Face Database. For more details, please see Langner et al. (2010). Presentation and validation of the Radboud Faces Database. *Cognition & Emotion*, 24(8), 1377–1388. doi: <https://doi.org/10.1080/0269993090348507>

color computer monitor and audio speakers. A digital video camera, located approximately 22 cm below the monitor, was connected to a computer and digital video recorder (DVR) in an adjoining control room. In the control room, an experimenter observed and recorded infants' looking times during each trial. Habit2 software (Oakes, Sperka, DeBolt, & Cantrell, 2019) was used to present the stimuli, record looking times, and calculate the habituation criteria (described below).

Procedure A habituation procedure called the “switch design” was used. After obtaining parental consent, infants were taken into the testing room and seated on their parent's lap in front of the computer monitor. Parents were instructed not to speak to their infant or point to the screen. Prior to each trial, an “attention-getter” (i.e., a green chiming, expanding circle) directed infants' attention to the monitor. Once the infant was looking at the screen, the experimenter initiated each trial and recorded infants' looking time. For a “look” to be counted, infants had to attend continuously for a minimum of 2 s. A trial ended when an infant either looked away for more than two continuous seconds or until the 30 s trial ended. Infants viewed one *Pre-test* trial, a maximum of 20 habituation trials, and three test trials (i.e., *Same*, *Switch*, *Post*).

The *Pre-test* trial was used to acclimate infants to the task. During this trial, infants viewed a facial configuration associated with surprise paired with the label “dax.” During the following habituation phase, infants viewed two different, alternating trials. In half of the habituation trials, infants viewed one of the target facial configurations (e.g., happy) paired with one of the labels (e.g., “blicket”). For the other half of the

trials, infants viewed the other target facial configuration (e.g., sad) paired with the other label (e.g., “toma”). Trials were randomized and presented in blocks, and infants could not view the same trial (e.g., happy-“blicket”) more than two times consecutively. The specific mappings of the labels and facial configurations were counterbalanced across participants. The habituation phase continued until infants' looking times across four consecutive trials decreased to 50% or more compared with their looking times during the longest four habituation trials (i.e., a sliding habituation window) (Oakes, 2010) or until all 20 habituation trials were presented.

Following habituation, infants viewed three test trials. The *Same* trial maintained one of the pairings viewed during habituation (e.g., happy-“blicket”). The *Switch* trial violated one of the habituation pairings (e.g., sad-“blicket”). Finally, the *Post* trial tested for infant fatigue and was identical to the *Pre-test* trial (i.e., surprise-“dax”). The presentation order of the *Same/Switch* trials and the pairing that “switched” were counterbalanced across participants. The *Post* trial was always presented last. After the testing session, parents completed the MacArthur Communicative Development Inventory Short Form (Fenson et al., 2007) to measure infants' (receptive) vocabulary.

Scoring Infants' looking was live-coded by a trained research assistant. The coder was blind to which stimuli the infant was viewing. A second coder, who was also display-blind, rescored 25% of the tapes ($n = 7$) offline. Reliability was excellent for duration of looking on each trial, $r = .99$, $p < .001$.

Results

For each experiment, all statistical tests were two-tailed and alpha was set at .05. Analyses were similar to previous research with this design (Casasola & Cohen, 2000; Werker et al., 2002). The dataset and analysis code for these studies (which includes participants excluded from final analyses) are available on OSF: https://osf.io/g6zej/?view_only=29d2804f0d6d4f4c996c411ec3f4b4e2.

Habituation Phase To ensure that infants demonstrated a significant decrease in looking time from habituation to test, a paired sample *t* test was conducted (Oakes, 2010). Infants' average-looking time to the longest four habituation trials ($M = 23.09$ s, $SD = 3.98$ s) was significantly longer than their looking time to the *Same* trial ($M = 10.42$ s, $SD = 5.48$ s), $t(28) = 11.78$, $p < .001$, $d = 2.19$.

Test Phase Infants' looking times to the three test trials (*Same/Switch/Post*) were analyzed in a within-subjects ANOVA. A significant main effect of test trials emerged, $F(2, 56) = 18.01$, $p < .001$, $\eta_p^2 = .39$ (Fig. 2). In contrast to our predictions, infants *did not* look significantly longer at the *Switch* trial ($M = 11.44$ s, $SD = 7.73$ s) compared with the *Same* trial ($M = 10.42$ s, $SD = 5.48$ s), $t(28) = .57$, $p > .25$, $d = .11$. This suggests that 14-month-olds did not map the novel labels to the facial configurations. To ensure that infants' failure to form the mapping was not due to fatigue at the end of the testing session, an additional comparison was conducted between the *Same* and *Post* trials. Infants looked significantly longer at the *Post* trial ($M = 19.77$ s, $SD = 8.26$ s), compared with the *Same* trial, $t(28) = 4.64$, $p < .001$, $d = .86$. Thus, infants were not fatigued by the end of the testing session.

Vocabulary The final analysis explored whether infants' receptive vocabulary was related to their ability to form the mappings. A difference score was computed between infants' looking time to the *Switch* trial and infants' looking time to the *Same* trial (for a similar analysis, see Werker et al., 2002). A regression analysis found that 14-month-olds with higher receptive vocabularies attended longer to the *Switch* trial compared with the *Same* trial, $F(1, 27) = 5.39$, $p = .028$, $R^2 = .17$.

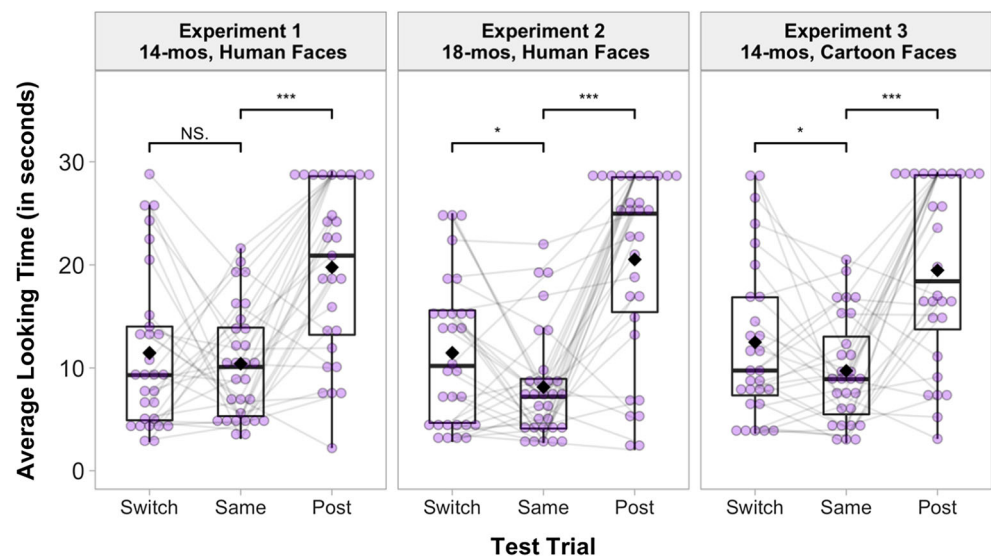
Experiment 2

Experiment 1 found that 14-month-olds did not map novel labels to facial configurations. This is in contrast to previous research in which 14-month-olds mapped novel labels to objects (Werker et al., 1998) and spatial relations (Casasola & Wilbourn, 2004). However, similar to previous research (Chan et al., 2011; Werker et al., 2002; Yoshida, Fennell, Swingley, & Werker, 2009), we found that 14-month-olds' vocabulary size related to their ability to form these mappings. Novice word learners may have difficulty in this task due to the cognitive complexity involved in associative word learning (Stager & Werker, 1997). Thus, it is possible that, with cognitive maturation and/or experience with language, infants could form these mappings. In Experiment 2, we predicted that older, 18-month-olds would map novel labels onto facial configurations.

Method

Participants A total of 62 18-month-olds were recruited and tested. The final sample consisted of 30 18-month-old infants (16 female, $M = 18.22$ months, $SD = .56$ months, range = 17.14 months–18.82 months). Participants were identified as

Fig. 2 Infants' average looking times to the test trials (in seconds), separated by Experiment. Gray lines between each test trial connect looking times for an individual participant. Comparisons between test trials, * $p \leq .05$, *** $p < .001$



White (86%, $n = 25$), Multi-Racial (10%, $n = 3$), Black (3%, $n = 1$), or not specified (3%, $n = 1$). Thirty-two additional infants participated in the study, but were excluded from final analyses for computer error ($n = 2$), failure to finish the experiment ($n = 7$), failure to meet the habituation criteria ($n = 7$), extreme looking to the *Same* test trial (longer than 25.20s; $n = 4$), inability to see eyes during the experiment ($n = 5$), fussiness or inattentiveness which led to difficulties with accurate coding ($n = 3$), or parental interference ($n = 2$). Infants “failed to finish” if either (a) the parent asked for the experiment to be stopped at any time or (b) the blind coder ended the experiment due to sustained infant crying over multiple trials. This attrition rate (51%) is similar to Experiment 1. All aspects of Experiment 2 were identical to Experiment 1 (Fig. 1).

Scoring A second coder re-scored 25% of the tapes ($n = 7$) offline. Reliability was excellent for duration of looking on each trial, $r = .99$, $p < .001$.

Results

Habituation Phase A paired sample t test showed that infants’ average looking time to the longest four habituation trials ($M = 25.23$ s, $SD = 3.87$ s) was significantly longer than their looking time to the *Same* trial ($M = 8.12$ s, $SD = 5.37$ s), $t(29) = 15.84$, $p < .001$, $d = 2.89$.

Test Phase A within-subject ANOVA revealed a significant main effect of test trials, $F(2, 58) = 26.39$, $p < .001$, $\eta_p^2 = .48$ (Fig. 2). In line with our hypotheses, infants looked significantly longer at the *Switch* trial ($M = 11.44$ s, $SD = 7.12$ s) compared with the *Same* trial ($M = 8.12$ s, $SD = 5.37$ s), $t(29) = 2.56$, $p = .016$, $d = .47$. This suggests that 18-month-olds mapped the novel labels to the facial configurations. An additional comparison found that infants also looked significantly longer at the *Post* trial ($M = 20.53$ s, $SD = 9.12$ s), compared to the *Same* trial, $t(27) = 8.12$, $p < .001$, $d = 5.37$, suggesting that infants were not fatigued by the end of the testing session.

Vocabulary A difference score was computed between infants’ looking time to the *Switch* trial and infants’ looking time to the *Same* trial. Unlike Experiment 1, 18-month-olds’ vocabulary size was not related to this difference score, $F(1, 28) = .05$, $p > .25$, $R^2 < .01$.

Experiment 3

Experiment 2 found that 18-month-olds were able to map novel labels to facial configurations. This developmental change between 14 and 18 months is similar to previous research on associative word learning with actions

(Casasola & Cohen, 2000; Chan et al., 2011). Thus, it appears that 14-month-olds have difficulty in this task due to the cognitive complexity involved in associative word learning (Stager & Werker, 1997). Perhaps for this reason, studies have also found that reducing the complexity of the task facilitates associative word learning for younger infants (Casasola & Cohen, 2000; Tsui et al., 2019; Werker et al., 2002; Yoshida et al., 2009). For example, 14-month-olds map labels to objects when the objects are presented in isolation against a black background (Werker et al., 1998), but not when the objects are embedded in a complex scene (Chan et al., 2011). Thus, Experiment 3 explored whether 14-month-olds could map novel labels onto *cartoon* facial configurations. We hypothesized that reducing the visual complexity of the stimuli would allow 14-month-olds to form these mappings.

Methods

Participants A total of 51 14-month-olds were recruited and tested. The final sample consisted of 28 14-month-old infants (15 female, $M = 14.10$ months, $SD = .27$ months, range = 13.52 months – 14.47 months). Participants were identified as White (68%, $n = 19$), Multi-Racial (18%, $n = 5$), Black (4%, $n = 1$), or did not specify (7%, $n = 2$). Twenty-three additional infants participated in the study but were excluded from final analyses due to failure to finish the experiment ($n = 10$), failure to meet the habituation criteria ($n = 4$), extreme looking to the *Same* test trial (longer than 22.83 s; $n = 4$), inability to see eyes during the experiment ($n = 2$), fussiness or inattentiveness which led to difficulties with accurate coding ($n = 4$), or parental interference ($n = 1$). This attrition rate (45%) is similar to Experiment 1.

Stimuli, Apparatus, and Procedure The visual stimuli were cartoon facial configurations created in Photoshop (Fig. 1). The stimuli were closely matched to the human facial configurations, including “human-like” eyes (i.e., visible sclera) as well as variable eyebrow positions and mouth shapes. All other aspects of Experiment 3 were identical to Experiment 1.

Scoring A second research assistant re-scored 25% of the tapes ($n = 7$) offline. Reliability was excellent for duration of looking on each trial, $r = .99$, $p < .001$.

Results

Habituation Phase A paired sample t test showed that average looking time to the longest four habituation trials ($M = 22.79$ s, $SD = 4.77$ s) was significantly longer than their looking time to the *Same* trial ($M = 9.72$ s, $SD = 5.19$ s), $t(27) = 13.00$, $p < .001$, $d = 2.46$.

Test Phase A within-subject ANOVA revealed a significant main effect of test trials, $F(2, 54) = 15.06$, $p < .001$, $\eta_p^2 = .36$ (Fig. 2). In line with our hypotheses, infants looked longer at the *Switch* trial ($M = 12.47$ s, $SD = 7.73$ s) compared with the *Same* trial ($M = 9.72$ s, $SD = 5.19$ s), $t(27) = 2.04$, $p = .051$, $d = .39$. This suggests that 14-month-olds mapped the novel labels onto the cartoon facial configurations. An additional comparison found that infants also looked significantly longer at the *Post* trial ($M = 19.46$ s, $SD = 8.92$ s), compared with the *Same* trial, $t(27) = 5.25$, $p < .001$, $d = .99$, suggesting that infants were not fatigued by the end of the testing session.

Vocabulary A difference score was computed between infants' looking time to the *Switch* trial and infants' looking time to the *Same* trial. Similar to Experiment 2, 14-month-olds' vocabulary size was not related to this difference score, $F(1, 26) = .33$, $p > .25$, $R^2 = .01$.

Discussion

The current studies are the first to examine whether infants can map novel labels to facial configurations. We found that 18-month-olds (Experiment 2), but not 14-month-olds (Experiment 1), map novel labels to *human* facial configurations. However, 14-month-olds formed these mappings with *cartoon* facial configurations (Experiment 3). This developmental pattern suggests that associative word learning with human facial configurations is of similar complexity to associative word learning with actions (e.g., pushing, pulling) (Casasola & Cohen, 2000; Chan et al., 2011) rather than associative word learning with objects (Werker et al., 1998). This is consistent with the active, dynamic displays of human emotions and suggests that the nature of emotion and action categories may be more abstract compared with object categories. We also found that more advanced word learners (as indexed by their receptive vocabulary and/or age) were better able to map novel labels to human facial configurations. More advanced word learners may (a) possess greater cognitive resources to process complex visual stimuli (i.e., *human* facial configurations) (Stager & Werker, 1997); (b) experience more social interactions with parents and caregivers, including greater exposure to others' emotions; and/or (c) excel at assimilating new words into their vocabularies (Widen, 2013). Since some degree of cognitive maturation, language experience, and/or emotional exposure is necessary to map labels to human facial configurations, this may partially explain why emotion labels emerge relatively late in children's vocabularies compared with other labels (e.g., "dog") (Fenson et al., 1994; Shablack et al., 2019).

Relatedly, it is interesting to note that vocabulary correlated with 14-month-old performance in Experiment 1, but not in Experiment 3. This discrepancy suggests that, due to the

complexity of Experiment 1, only infants with more advanced vocabularies could form the associative links with human facial configurations. However, by reducing the task demands in Experiment 3 (i.e., using less complex visual stimuli), infants with lower language abilities could form the associative links with cartoon facial configurations. These cartoon faces were schematic symbols that, compared with human emotion categories, are more homogenous, less abstract, and relatively impoverished in terms of facial features (e.g., noses, ears, hair). Thus, although 14-month-olds could form associative links with labels and these facial configurations, cartoon faces have relatively low ecological validity. Instead, the capacity to associate labels with abstract, heterogeneous, and ecologically valid human facial configurations does not appear to fully emerge until 18 months of age.

Other explanations for these findings should be addressed. First, it is possible that since parents could see the screen, they involuntarily provided cues to their infants. This seems unlikely since (a) we did not include parents in the final analyses who pointed at the screen or told their infants to look at the screen (i.e., exclusions for parental interference) and (b) infants' looking systematically differed across the three experiments (i.e., it is unlikely that parents provided cues in Experiment 2 and 3 but not Experiment 1). Another possibility is that 14-month-olds in Experiment 1 had already associated labels with the facial configurations and, thus, would not map an additional label to these stimuli (i.e., "mutual exclusivity"). While this explanation is possible, it is unlikely for two reasons. First, while mutual exclusivity constraints on word learning exist as early as 15 months of age (e.g., Markman, Wasow, & Hansen, 2003), it is unlikely that 14-month-olds, but not older, 18-month-olds, were constrained by this principle. Second, few 14-month-olds (or 18-month-olds) in our sample have the words "happy," "sad," "smile," or "frown" in their receptive or productive vocabularies (see the Supplementary Materials).

These studies found that infants can learn arbitrary associations between labels and facial configurations without rich contextual or social support. Additional future research is needed to explore how infants form these associations in more naturalistic settings. Specifically, while the presence of social partner may help word learning (Werker et al., 1998), additional contextual information may further complicate this process (Chan et al., 2011). Similarly, it is possible that with increased task complexity (i.e., using human facial configurations associated with within-valence emotions), 18-month-olds would also struggle to form these associations (Ruba & Repacholi, 2019). In order to assess the process by which infants are learning "emotion" labels, specifically, we tested facial configurations associated with different emotions. However, we believe that 18-month-olds could also succeed at this type of task if "novel" facial configurations were used that do not have clear emotion associations (DiGirolamo &

Russell, 2017). In this way, it is unclear whether infants mapped labels to emotion states conveyed by the pictures (i.e., “happiness”) or to the perceptual patterns of facial features (i.e., a smile). It is also an open question as to whether these pairings persist over time (i.e., whether infants learned the words), as could be assessed using alternative experimental paradigms. One potential limitation is the high attrition rate in these studies (38–51%), which was driven by a failure to habituate and infant fussiness. While this attrition rate is not unprecedented for the switch design, it may indicate that this paradigm is too cognitively taxing for infants at this age. Thus, future research may consider alternate paradigms to measure associative word learning with 14- and 18-month-olds.

These findings also provide insights to current debates in affective science regarding the role of language in emotion concept development. While some have hypothesized that learning emotion labels is necessary for infants to develop concepts of within-valence emotions (e.g., anger v. disgust) (Hoemann et al., 2019; Shablack, Stein, & Lindquist, 2020), others have argued that infants begin to develop these concepts prior to learning emotion labels (Ruba & Repacholi, 2019, 2020). Emerging empirical evidence appears to support the latter interpretation, showing that infants may have some conceptual knowledge of within-valence emotions prior to 18 months of age (Ruba, Meltzoff, & Repacholi, 2019, 2020a; Wu, Muentener, & Schulz, 2017). Considered alongside the current studies, this suggests that infants begin to develop within-valence emotion concepts *before* they develop the capacity to map labels to human facial configurations at 18 months. Although the current studies tested facial configurations associated with between-valence emotions (i.e., happy v. sad), associating labels with within-valence emotions is likely an even more difficult task (Ruba & Repacholi, 2019; Widen, 2013). Thus, it is unlikely that infants younger than 18 months can map labels to within-valence facial configurations. While the process of learning emotion labels may influence further development of these emotion concepts (Ruba et al., 2020b), this remains a relatively unstudied area of affective science. In sum, these studies provide the first empirical evidence that, by 18 months of age, infants can map labels to human facial configurations quickly without explicit instruction. This process likely has a critical, constructive influence on emotion concept development across the lifespan.

Additional Information

Acknowledgments The authors thank the Wilbourn Infant Lab at Duke for assistance with data collection.

Funding ALR was supported by an Emotion Research Training Grant from NIMH (T32-MH018931).

Data Availability The dataset and analysis code for these studies are available on OSF: https://osf.io/g6zej/?view_only=29d2804f0d6d4f4c996c411ec3f4b4e2.

Conflict of Interest On behalf of all authors, the corresponding author states that there is no conflict of interest.

Ethical Approval The study was performed to ethical standards as laid down in the 1964 Declaration of Helsinki and conducted with approval of the institutional review board at Duke University (Approval Number: B0181, Protocol Title: “Early Language Learning: Labels for Emotional Expressions”).

Informed Consent All parents provided informed consent for their infants to participate in the study.

References

- Casasola, M., & Cohen, L. B. (2000). Infants’ association of linguistic labels with causal actions. *Developmental Psychology*, *36*(2), 155–168. <https://doi.org/10.1037/0012-1649.36.2.155>.
- Casasola, M., & Wilbourn, M. P. (2004). Fourteen-month-old infants form novel word-spatial relation associations. *Infancy*, *6*(3), 385–396. https://doi.org/10.1207/s15327078in0603_4.
- Chan, C. C. Y., Tardif, T., Chen, J., Pulverman, R. B., Zhu, L., & Meng, X. (2011). English- and Chinese-learning infants map novel labels to objects and actions differently. *Developmental Psychology*, *47*(5), 1459–1471. <https://doi.org/10.1037/a0024049>.
- DiGirolamo, M. A., & Russell, J. A. (2017). The emotion seen in a face can be a methodological artifact: The process of elimination hypothesis. *Emotion*, *17*(3), 538–546. <https://doi.org/10.1037/emo0000247>.
- Fennell, C. T., & Werker, J. F. (2003). Early word learners’ ability to access phonetic detail in well-known words. *Language and Speech*, *46*(2–3), 245–264. <https://doi.org/10.1177/00238309030460020901>.
- Fenson, L., Dale, P. S., Reznick, J. S., Bates, E., Thal, D. J., Pethick, S. J., Tomasello, M., Mervis, C. B., & Stiles, J. (1994). Variability in early communicative development. *Monographs of the Society for Research in Child Development*, *59*(5), i. <https://doi.org/10.2307/1166093>.
- Fenson, L., Marchman, V. A., Thal, D. J., Dale, P. S., Reznick, J. S., & Bates, E. (2007). *MacArthur-Bates communicative development inventories*. https://www.uh.edu/class/psychology/dcbn/research/cognitive-development/_docs/mcdigestures.pdf
- Hoemann, K., Xu, F., & Barrett, L. F. (2019). Emotion words, emotion concepts, and emotional development in children: A constructionist hypothesis. *Developmental Psychology*, *55*(9), 1830–1849. <https://doi.org/10.1037/dev0000686>.
- Izard, C. E., Fine, S., Schultz, D., Mostow, A., Ackerman, B. P., & Youngstrom, E. A. (2001). Emotion knowledge as a predictor of social behavior and academic competence in children at risk. *Psychological Science*, *12*(1), 18–23. <https://doi.org/10.1111/1467-9280.00304>.
- Langner, O., Dotsch, R., Bijlstra, G., Wigboldus, D. H. J., Hawk, S. T., & van Knippenberg, A. (2010). Presentation and validation of the radboud faces database. *Cognition and Emotion*, *24*(8), 1377–1388. <https://doi.org/10.1080/02699930903485076>.
- Markman, E. M., Wasow, J. L., & Hansen, M. B. (2003). Use of the mutual exclusivity assumption by young word learners. *Cognitive Psychology*, *47*(3), 241–275. [https://doi.org/10.1016/S0010-0285\(03\)00034-3](https://doi.org/10.1016/S0010-0285(03)00034-3).
- Nelson, N. L., & Russell, J. A. (2016). 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>.
- Oakes, L. M. (2010). Using habituation of looking time to assess mental processes in infancy. *Journal of Cognition and Development*, 11(3), 255–268. <https://doi.org/10.1080/15248371003699977>.
- Oakes, L. M., Sperka, D., DeBolt, M. C., & Cantrell, L. M. (2019). Habit2: A stand-alone software solution for presenting stimuli and recording infant looking times in order to study infant development. *Behavior Research Methods*, 51, 1943–1952. <https://doi.org/10.3758/s13428-019-01244-y>.
- Pons, F., Harris, P. L., & de Rosnay, M. (2004). Emotion comprehension between 3 and 11 years: Developmental periods and hierarchical organization. *European Journal of Developmental Psychology*, 1(2), 127–152. <https://doi.org/10.1080/17405620344000022>.
- Ruba, A. L., Meltzoff, A. N., & Repacholi, B. M. (2019). How do you feel? Preverbal infants match negative emotions to events. *Developmental Psychology*, 55(6), 1138–1149. <https://doi.org/10.1037/dev0000711>.
- Ruba, A. L., Meltzoff, A. N., & Repacholi, B. M. (2020a). The development of negative event-emotion matching in infancy: Implications for theories in affective science. *Affective Science*, 1(1), 4–19. <https://doi.org/10.1007/s42761-020-00005-x>.
- Ruba, A. L., Meltzoff, A. N., & Repacholi, B. M. (2020b). Superordinate categorization of negative facial expressions in infancy: The influence of labels. *Developmental Psychology*, 56(4), 671–685. <https://doi.org/10.1037/dev0000892>.
- Ruba, A. L., & Repacholi, B. M. (2019). Do preverbal infants understand discrete facial expressions of emotion? *Emotion Review*, 175407391987109. <https://doi.org/10.1177/1754073919871098>.
- Ruba, A. L., & Repacholi, B. M. (2020). Beyond language in infant emotion concept development. *Emotion Review*.
- Shablack, H., Becker, M., & Lindquist, K. A. (2019). How do children learn novel emotion words? A study of emotion concept acquisition in preschoolers. *Journal of Experimental Psychology: General*, 149, 1537–1553. <https://doi.org/10.1037/xge0000727>.
- Shablack, H., Stein, A. G., & Lindquist, K. A. (2020). Comment: A role of language in infant emotion concept acquisition. *Emotion Review*, 175407391989729. <https://doi.org/10.1177/1754073919897297>.
- Stager, C. L., & Werker, J. F. (1997). Infants listen for more phonetic detail in speech perception than in word-learning tasks. *Nature*, 388(6640), 381–382. <https://doi.org/10.1038/41102>.
- Tsui, A. S. M., Byers-Heinlein, K., & Fennell, C. T. (2019). Associative word learning in infancy: A meta-analysis of the switch task. *Developmental Psychology*, 55(5), 934–950. <https://doi.org/10.1037/dev0000699>.
- Werker, J. F., Cohen, L. B., Lloyd, V. L., Casasola, M., & Stager, C. L. (1998). Acquisition of word–object associations by 14-month-old infants. *Developmental Psychology*, 34(6), 1289–1309. <https://doi.org/10.1037/0012-1649.34.6.1289>.
- Werker, J. F., Fennell, C. T., Corcoran, K. M., & Stager, C. L. (2002). Infants’ ability to learn phonetically similar words: Effects of age and vocabulary size. *Infancy*, 3(1), 1–30. https://doi.org/10.1207/S15327078IN0301_1.
- 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>.
- 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>.
- Yoshida, K. A., Fennell, C. T., Swingley, D., & Werker, J. F. (2009). Fourteen-month-old infants learn similar-sounding words. *Developmental Science*, 12(3), 412–418. <https://doi.org/10.1111/j.1467-7687.2008.00789.x>.