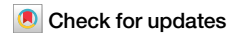




Infants' facial input from 6 to 17 months rarely includes hypothesized expressive emotion prototypes



Katie Hoemann¹ , Jeffrey M. Girard¹ , Marissa Ogren², Lauren Leotti², Lisa Feldman Barrett^{3,4,5} , Lisa M. Oakes^{6,7} & Vanessa LoBue²

The facial configurations that infants encounter in typical interactions with caregivers and family members form a basis for their understanding of emotion. Studies to date have described or quantified these natural exposures to facial information, but descriptions of the emotional content of that input are still needed. We accomplished this by leveraging an existing, longitudinal corpus of video-recordings of 6- to 17-month-old infants' interactions with close others. We found that infants were generally exposed to far more partial than full faces, and that exposure to full faces, while low, increased over the time periods observed. Exposure to facial configurations hypothesized to serve as expressive prototypes for categories of emotion (e.g., HAPPINESS, ANGER) was extremely rare. When hypothesized expressive prototypes did occur, it was often in unexpected contexts, suggesting that they did not necessarily serve as a cue to the presumed emotional meaning. Lastly, descriptive analyses of the facial movement data revealed that infants' facial input was dominated by a few, mostly simple facial configurations that did not correspond fully or uniquely to hypothesized category prototypes. These results extend evidence that emotions are expressed heterogeneously in daily life, such that infants must learn much more than one-to-one mappings between facial configurations and context/behavior to acquire expressive vocabularies for emotion.

How do infants acquire an understanding of emotions and the ability to infer them in others? For decades, researchers have attempted to answer this question by presenting infants with static images of adults making facial configurations thought to have one-to-one relations with particular categories of emotion, such as scowling faces as expressions of anger. This work led to the conclusion that discrimination, categorization, and identification of emotional facial expressions begins in infancy, and proceeds at least through early childhood (for a review, see ref. 1). Importantly, a series of meta-analyses has shown that such one-to-one relations do not exist in the real world^{2,3} (for discussion, see refs. 4,5). Research with both adults³ and children^{2,6,7} suggests instead that we express emotions with considerably more variability than the typical posed configurations represent⁴. For example, meta-analytic evidence indicates that adults only scowl about 35% of the time when angry, meaning that 65% of the time they express anger in other ways³. Moreover, only half the scowls that people make are

expressions of anger⁸. People furrow their brows and purse their lips at times when they are not angry, such as when they are confused, concentrating, or looking toward the sun. This suggests that infants may not be exposed to the types of diagnostic information from which we have long assumed they learn emotions. However, there are only a handful of studies to date documenting the actual facial input infants receive from their everyday environments, and these studies have not captured the *emotional content* of that input. The goal of the current investigation is to characterize the facial configurations that infants encounter in their daily lives, which serve as the curriculum for their developing emotion understanding.

Most work on infants' acquisition of emotion understanding has been based on the assumption that categories of emotion—like ANGER, FEAR, SADNESS, and HAPPINESS—have exemplars that share a representative set of features^{9–11} (as discussed in refs. 4,12). These proposed features include configurations of facial movements that are considered to be reliable and

¹Department of Psychology, University of Kansas, Lawrence, KS, USA. ²Department of Psychology, Rutgers University, Newark, NJ, USA. ³Department of Psychology, Northeastern University, Boston, MA, USA. ⁴Psychiatric Neuroimaging Division, Department of Psychiatry, Massachusetts General Hospital, Charlestown, MA, USA. ⁵Athinoula A. Martinos Center for Biomedical Imaging, Department of Radiology, Massachusetts General Hospital, Charlestown, MA, USA. ⁶Center for Mind and Brain, University of California at Davis, Davis, CA, USA. ⁷Department of Psychology, University of California at Davis, Davis, CA, USA.

 e-mail: khoemann@gmail.com

specific cues for the presence of a given emotion in most circumstances. For example, scowls are thought to indicate anger with sufficient consistency, and not to indicate sadness, confusion, etc. This *prototype view* of emotion has led to the design of studies in which certain posed, often exaggerated facial configurations are taken as the standards for specific emotion categories^{13–24}. Although this literature provides an important baseline for the developmental trajectory of early emotion understanding, the results of these studies are only meaningful in representing how infants acquire real-world emotion knowledge if these facial configurations align with the distribution of input infants naturally and routinely encounter.

It is possible that adults interacting with infants make prototypical facial configurations and use them to mark instances of presumably corresponding emotions (e.g., scowling for ANGER). Adults in many cultures indeed use exaggerated, simplified speech, or child-directed speech (i.e., ‘motherese’), when talking to infants^{25,26} and adults from middle-class North American families use exaggerated, simplified motions (i.e., ‘motionese’) when demonstrating actions to infants²⁷. These modifications of speech and action may facilitate infants’ learning of that input. Thus, adults may also use more exaggerated, stereotypical facial configurations when interacting with infants to intentionally grab infants’ attention or tag associations with actions and words. For example, caregivers may pair dramatic expressive facial movements with particular contexts or behaviors to signal (un)desirable outcomes (e.g., intense scowling after the infant throws their food on the floor) or with specific labels to teach emotion categories (e.g., ‘happy!’ paired with an expansive smile).

As reviewed above, however, growing empirical evidence suggests that the input for infants’ learning of emotions consists of much more variable exemplars^{2,3,6–8,28}. The facial configurations used to express an emotion also vary as a function of context (see Ref. 4 and references therein), such that those hypothesized as category prototypes might not be the most common type of information infants receive about emotions in the real world. Thus, instead of the one-to-one relation between facial configurations and particular categories of emotion often assumed in past research, there may be many-to-many mappings between facial movements and behaviors/labels²⁹ (for discussion, see refs. 4,5,30,31). Caregivers might frown at thrown food because they are confused or make beleaguered smiles when trying to cheer a crying infant even as they are annoyed, and these facial configurations might be accompanied by seemingly incongruous labels (e.g., ‘why are you angry?’, ‘don’t be sad’) or no emotion labels at all. This diverse array of situated behaviors paints a very different picture of the learning problem that infants face in developing emotion understanding.

Yet research that directly documents the variation in facial configurations infants actually see in their everyday interactions with other people is still needed. To fully comprehend their learning problem, we must first characterize the distribution of input that infants use to build expectations about social interactions, including inferring the emotions of others, and planning their behavior adaptively. The value of input for gaining insight into developmental processes is increasingly recognized, as input can change over the course of development³² and these changes can create new contexts for learning³³. Although some research has examined infants’ everyday exposure to human faces, and how that exposure changes across development^{34,35}, the kinds of facial configurations infants experience in typical adult-infant interactions is virtually unknown. Here, we took a step toward meeting this need by analyzing a corpus of naturalistic interactions between infants and close social others. The Study of Environmental Effects on Developing Linguistic Skills (SEEDLingS³⁶) is an existing corpus of 44 infants, including hour-long video recordings taken in their homes once a month from 6 to 17 months of age. Infants wore a hat with two cameras—one camera oriented slightly up and the other slightly down, to best capture the infant’s field of view—giving us access to their visual emotional input. By sampling each infant at four ages (6–8 months, 9–11 months, 12–14 months, and 15–17 months), we were also able to track changes in exposure to facial input over time.

We began by assessing infants’ exposure to faces and facial configurations. Building on prior studies^{34,35,37}, we first quantified how often faces

were available in infants’ field of view. Videos were annotated for the presence of a full or partial face. Full faces were further annotated for facial muscle movements by an expert coder trained in the Facial Action Coding System (FACS³⁸). We assessed the percentage of time each infant was exposed to partial and/or full faces and how this varied over the course of the study. We then examined whether the observed facial movements mapped onto a large set of expressions that have been hypothesized as prototypes for individual emotion categories^{11,39–42}. Finally, we catalogued the actual configurations presented to infants to characterize their facial input. Together, these analyses provide an initial description of the facial information available to infants in this age range.

Methods

Data Set

In the current project, we applied new coding and analytical procedures to an existing video data set—The Study of Environmental Effects on Developing Linguistic Skills (SEEDLingS³⁶). Original collection of the SEEDLingS dataset was approved through the University of Rochester Subjects Review Board under RSRB No. 51231 and was performed in accordance with the Declaration of Helsinki. As detailed below, informed consent was obtained from all adult participants included in the study. Secondary data analysis was approved by the Rutgers University Institutional Review Board (IRB) under Pro2020000399 and reviewed (and deemed exempt) by the University of Kansas IRB under STUDY00151856. SEEDLingS data are available via Databrary; video data were accessed between January 2021 and February 2023. Our coding and analyses were not preregistered.

The data set contains hour-long videos of 44 infants ($n = 21$ female infants, $n = 23$ male infants; $n = 42$ white, $n = 1$ Black, $n = 1$ multiracial) engaging in natural interactions with people in their home environment. All participating families lived within a one-hour drive of Rochester, NY. The sample population was highly homogenous, with 83% of participants identifying as non-Hispanic and white, and 75% of mothers holding at least a bachelor’s degree. All families were two-parent households. There was no requirement that any specific family member was or was not present for the recordings, and so family composition varied from visit to visit⁴³. For more information about infants’ exposure to unique faces in the home sessions, see Supplementary information A.

Parents were initially consented for the entire yearlong study through a process overseen by the original university IRB. Additionally, for each recording, they consented to a level of sharing that they felt comfortable with; only recordings that parents gave explicit permission for sharing with researchers on platforms were shared with Databrary, where the present authors accessed them. The sharing occurred after recordings were scrubbed of personal information (in the audio stream, video stream, or both, e.g., full names, addresses, etc.); faces were not blurred. Participants were compensated \$340 for the yearlong study. See ref. 44 for further details.

Video data included views from one tripod (located centrally in the room where the infant was physically located during the session) and two cameras affixed to a hat or headband placed on the infant’s head. These head-mounted cameras provided information about what was within the infant’s field of view: one was oriented slightly up and the other slightly down, with visual field capture verified via Bluetooth connection during setup. Both head-mounted cameras were used for subsequent video coding, as faces could be available in either. At later time points (largely corresponding with the 12- and 15-month age windows described next), primary caregivers also wore a head-mounted camera attached to a hat.

For SEEDLingS, the same infants were repeatedly filmed once per month from 6 months to 17 months (a maximum of 12 total time points). However, to maintain a manageable coding burden for the current project, we assessed video data in four windows, with each infant providing video data a maximum of one time per window (6–8, 9–11, 12–14, and 15–17 months). If there was usable video data (defined as head-mounted video camera data available for at least 45 min out of the hour-long session) for a given infant at the youngest time point in each age window (e.g., 6 months), then the video from that age was used. If not, then the next age

was assessed (e.g., 7 months). If there was no usable data for a given infant at either of the youngest two ages in a given window, the oldest age was used (e.g., 8 months). If no usable video data was provided for an entire window (e.g., 6–8 months), then this time point was considered missing for that infant. The number of videos included in the final data set at each age are as follows: 6 months, $n = 36$; 7 months, $n = 7$; 8 months, $n = 1$; 9 months, $n = 30$; 10 months, $n = 7$; 11 months, $n = 4$; 12 months, $n = 16$; 13 months, $n = 5$; 14 months, $n = 5$; 15 months, $n = 16$; 16 months, $n = 9$; 17 months, $n = 0$. By window, this is: 6–8 months, $n = 44$; 9–11 months, $n = 41$; 12–14 months, $n = 26$; 15–17 months, $n = 25$. In total, 136 h-long videos were coded and analyzed for the current project. In the final data set, 19 infants provided usable data for all four age windows, 11 infants for three windows, 13 infants for two windows, and 1 infant for only one window. At each time point, multiple aspects of the video data were coded in stages, as described below.

Face availability coding

Video data were coded continuously by human coders using Datavyu⁴⁵. To determine the duration of face availability to infants, the entire duration of each session was coded as either having a face available or unavailable. A face was coded as available when any part of the mouth, nose, or eyes (including eyebrows) was visible in either of the infant head-mounted cameras. Otherwise, the time was coded as “face unavailable.” Static images of faces (e.g., photos) or faces on TV were not coded. Videos varied in duration across participants and time points, but they were all approximately one hour in length. The entire recorded session was coded, excluding any time the experimenters were visible or audible at the start or end of a session, and any time the cameras were turned off or removed. To ensure that coding was consistent across the team of research assistants, all videos were coded by a primary coder, and a second coder performed reliability coding on 20% of videos. Reliability between coders was high ($ICC = 0.94$). Data from the primary coder were used for all analyses.

In addition to looking at the availability of any facial information, we examined the amount of time that at least one *entire face* was visible on camera. Building on the face availability codes described above, a face was coded as *fully visible* if all of the following features were visible: both complete eyebrows and eyes (including the lower eyelid), the forehead, the nose (including the bridge), lips, chin, and both nasolabial furrows (lines in the face that run roughly from the nose to the outer corner of the lips). To ensure that fully visible faces were close enough to the camera for potential Facial Action Coding System (FACS) coding, if eyes were open, sclera (whites of the eyes) had to also be visible. If eyes were closed, the top eyelid had to be fully visible. If these criteria were not met, the face was coded as *partially visible*.

Facial configuration coding

The Facial Action Coding System (FACS³⁸) was used to determine the objective facial configurations that were within infants’ field of view of infants based on observable facial movement. A trained, FACS-certified coder (author M.O.) completed all FACS coding in Datavyu on a frame-by-frame basis for all segments when a face had been identified as fully visible. FACS codes were used to indicate which facial muscles were moving on the visible face within the specified time window. Oftentimes, multiple distinct FACS codes were documented within a single segment, as the individual’s facial configuration changed while still fully visible to the infant. All clearly visible faces were FACS coded unless a substantive portion was unclear due to lighting or motion or otherwise occluded. For instance, the head-mounted (“hat”) camera worn by caregivers at later time points covered the forehead, limiting visibility of corresponding facial muscle movements.

Action units (AUs) 1–39 were used in the current project, describing common muscle movements in the upper and lower face. Codes for blinks (AU 45) were omitted, although squinting was still coded as AU 6 and/or 7 as relevant. Further, codes for speech (AU 50), lips parted (AU 25), and jaw dropped (AU 26) were not coded in isolation. These FACS codes were only coded if they occurred in conjunction with another facial muscle movement

(e.g., smiling while talking would be coded as 12 + 50). This decision was made to substantially reduce the coding burden, as these particular FACS codes were deemed to carry limited emotional information in isolation. Further, head and eye positions were not coded; neither were hand movements or body posture. Laterality and intensity of movements were coded for all applicable movements but were omitted from the present analyses. If no observable facial muscle movements occurred within a given segment (i.e., a truly neutral face was present), AU 0 was documented. Subsequent analyses only included full faces that were not occluded (i.e., that were coded with any AU, including AU 0).

We examined the proportion of video frames with full faces that could be FACS coded per age window, grouping these frames into those coded with AU 0 (neutral) versus any other AU, or occluded due to the caregiver hat camera versus another source (e.g., shadow). We found that 62% and 70% of full-face frames could not be coded in the 12- and 15-month windows, respectively, due to the caregiver-worn camera (compared to 0% and 3% in the 6- and 9-month windows; see Fig. S5 in Supplementary Information B). Many of the coded movements in these windows came from a secondary source (e.g., another child), compromising the comparability of these later windows with the earlier time points.

Hypothesized expressive prototypes

We consulted both recent and core inventories in the science of emotion to compile a comprehensive set of the facial muscle movements that are hypothesized to serve as expressive prototypes^{11,39–42}. These hypothesized expressive prototypes can refer explicitly to configurations—that is, sets of simultaneous facial movements (e.g., ref. 42)—or to probabilistic combinations—that is, movements that are more likely to occur in tandem (e.g., ref. 40). We excluded prototypes that are thought to be compounds of multiple emotion categories (e.g., “happily surprised”⁴¹) and prototypes consisting solely of FACS codes that were not coded in the current project (e.g., BOREDOM as AU 43, INTEREST as AU 85⁴⁰). The remaining prototypes were included regardless of whether they contained additional, non-AU movements (e.g., EMBARRASSMENT includes a face touch¹¹, which was not coded). From ref. 40, we used the international core patterns (ICPs), discovered as AUs occurring at above-chance frequency for a given emotion category across cultural contexts. All possible variants were considered (e.g., HAPPINESS could be indicated by the combination of AUs 12 + 25, as well as 6 + 12 + 25⁴¹). Hypothesized expressive prototypes were categorized as indicated by the original authors, with the exceptions of “joy” and “happiness”⁴², which were considered synonyms, and “desire food” and “desire sex”⁴⁰, which were merged into a single “desire” category. The above decisions resulted in a total of 264 unique and assessable expressive prototypes from 21 emotion categories: AMUSEMENT, ANGER, AWE, CONFUSION, CONTEMPT, CONTENTMENT, COYNESS, DESIRE, DISGUST, EMBARRASSMENT, FEAR, HAPPINESS, INTEREST, PAIN, PRIDE, RELIEF, SADNESS, SHAME, SURPRISE, SYMPATHY, and TRIUMPH. Note that, in several instances, the same facial configuration was hypothesized to be an expressive prototype of multiple, distinct emotion categories (e.g., AU 4 + 7 as ANGER and CONFUSION).

Exploratory examination of context

In addition to identifying a comprehensive set of hypothesized expressive prototypes and their corresponding emotion categories, we performed an exploratory examination of the context in which a subset of prototypes occurred for 5 of the 6 emotion categories hypothesized to be “basic” and universal: ANGER, DISGUST, FEAR, HAPPINESS, and SADNESS. This subset consisted of the 59 prototypes for these emotion categories as stipulated in the FACS Investigator’s Guide³⁹. Prototypes for SURPRISE were not examined because of its ambiguous valence.

First, a FACS-certified coder (author M.O.) identified full faces containing combinations of FACS codes that were exact matches for these hypothesized expressive prototypes (i.e., no missing or additional AUs). Next, trained research assistants reviewed the 10-s window of video recording surrounding each exact match (i.e., 5 s before and after) and determined the general context of the interaction (e.g., playing with toys

together, mealtime, and diaper/clothing changes), as well as whether the context was positive or negative in valence, based on whether the infant (or sibling present) was fussing/crying or experiencing a potential threat (e.g., showing signs of illness, choking, or risk of injury). For each window, the research assistant identified whether the context matched the valence of the hypothesized emotion category (i.e., positive or negative). In other words, hypothesized expressions of HAPPINESS in positive contexts were coded as matches, as were hypothesized expressions of ANGER, DISGUST, FEAR, and SADNESS in negative contexts. For example, if a child was crying and the parent demonstrated a combination of FACS codes consistent with a SADNESS prototype, the expression-context pairing was coded as a match. However, if a parent displayed FACS codes consistent with a FEAR prototype but the child was playing and there was no obvious threat, the expression-context pairing was coded as a mismatch.

Data pre-processing

Datavyu files for facial availability and facial configuration (FACS codes) were combined across all time points and infants in long format, with separate variables capturing infants' age at the video (e.g., 7 months) and corresponding age window (e.g., 6–8 months). The facial availability and facial configuration data were originally coded in terms of onset and offset timestamps. We used these timestamps to calculate duration (in seconds), percentages of video frames (i.e., session time) containing faces or configurations, and percentages of full-face frames containing configurations, as specified below.

Data analysis

Facial presence. Hierarchical Bayesian models (HBMs⁴⁶) were used to analyze the percentage of each video's frames that contained partial and full faces. The HBM framework combines generalized linear mixed modeling with Bayesian estimation and offers many theoretical and practical advantages over frequentist approaches⁴⁷. These models allowed us to flexibly account for several challenges of the data and ensure that the model assumptions matched the data. First, the fact that multiple videos were nested within each infant was accounted for using random intercepts and slopes per infant (as is typical in a mixed-effects model). Second, the fact that the outcome variable was a percentage (and therefore had non-normal residuals) was accounted for using the beta distributional family⁴⁸. Third, the fact that some observations were missing was accounted for using imputation during model fitting⁴⁹. Fourth, the relatively small size of the dataset was accounted for using regularizing priors, enabling the inclusion of the recommended maximal random effects structure⁵⁰. Bayesian diagnostic plots were used to confirm that data assumptions of the models were met.

We fit the same two types of HBM to the partial and full face presence variables, resulting in four total models. The first model for each was an unconditional or “null” model that included no predictors (see Supplementary information C.i); these were used to calculate the Bayesian equivalent of intraclass correlation coefficients (see Supplementary information C.ii), which indicate how much of the variance in the outcome variable was between (as opposed to within) infants. The second model for each was a conditional model that included the age window of each video (coded numerically as 0 = 6–8 months, 1 = 9–11 months, 2 = 12–14 months, or 3 = 15–17 months) as a linear predictor with random slopes per infant (see Supplementary information C.iii); these growth curves were used to test whether face availability increased over time within the average infant. As a sensitivity analysis, we also fit models that relaxed the assumption of linear growth by using monotonic effects⁵¹ for the age window predictor; the same pattern of significant results was found for these models, increasing our confidence in their accuracy (see Supplementary information C.iv). Full details of the analyses, including R code (version 4.5.1) and all output, are provided via our OSF repository. In summary, all models were fit using the brms R package⁵², normal priors centered on zero for slope parameters, default weakly informative priors for all other parameters, and four Markov chains (each with 9000 warmup iterations and 1000 inference iterations).

Hypothesized expressive prototypes. We examined whether the configurations of facial movements that occurred in full faces mapped onto one of the hypothesized expressive prototypes. We calculated the total duration of exposure to each of the 269 assessable prototypes and divided it by the total valid session time and the total valid full-face time, pooled across infants within each set of age windows (Fig. 2). For a coded (i.e., observed) facial configuration to be classified as a hypothesized expressive prototype, it had to exactly match one of the configurations listed in our source documents (i.e., all hypothesized AUs had to be present and no additional AUs could be present). The only exceptions to this rule were when a hypothesized expressive prototype included action descriptions (ADs) that are either not included in FACS (e.g., “fist in air” for TRIUMPH or “face touch” for EMBARRASSMENT) or AUs not included in our coding scheme (i.e., the set of codes that our FACS coder annotated). The relevant codes that were omitted from our scheme include: eye closure (AU 43), head positions (ADs 51–58), and eye positions (ADs 61–66). In a total of 46 cases (25 of which were for SADNESS), we removed the uncoded actions from the hypothesized expressive prototype before assessing its match with the coded configuration. These removals make it more likely for matches to occur and thus may slightly inflate the observed prevalence of the affected hypothesized expressive prototypes. Facial configurations could match more than one hypothesized expressive prototype; emotion counts were not additive. If multiple faces were visible at the same time, their exposure durations were summed and then weighted by the total session (or full-face) time per infant.

Considering the reduced comparability of data in the later time points, we analyzed 6- and 9-month windows separately from 12- and 15-month windows. For this same reason, we were unable to test whether infants' exposure to hypothesized prototypes changed with across the age range sampled by SEEDLingS.

Observed facial configurations. We used a data-driven approach to characterize the facial configurations infants were exposed to. This analysis included only data from the 6- and 9-month age windows due to data missingness at 12- and 15-months created by the caregiver-worn cameras. We treated each coded FACS event as one momentary configuration of a single face; because recordings often contained more than one fully visible face, coded separately by identity, configurations from co-visible faces were treated as distinct and simultaneous rather than merged. A configuration was defined as the exact set of coded AUs present during an event (excluding AU 50), such that configurations differing by a single AU were treated as distinct. After removing events with missing identity or zero duration, 11,077 configuration occurrences remained. Each distinct configuration was quantified by its number of events, the number of distinct infants and distinct individuals in whose input it appeared, its total visible duration, and the percentage of session and full-face time it occupied.

Additionally, we compared each observed configuration against the hypothesized expressive prototypes described above, reduced to our coded AU set as before. Following prior work^{8,40}, we calculated a Dice coefficient for every configuration-prototype combination and selected the best match. We interpreted the resulting distribution of match scores based on previously-established correspondence thresholds^{8,53}: none <0.2; weak 0.2–0.4; moderate 0.4–0.7; strong ≥0.7. Hypothesized expressive prototypes that contained no coded AUs were unassessable and excluded from scoring (5 or 1.8% of prototypes; see Table S3 in Supplementary information D); those missing some AUs were partially assessable and flagged separately for this reason.

Results

Facial presence

The first set of models revealed that the percentage of each video's duration that contained a partial face ranged from 3% to 65% ($M = 32\%$, $SD = 13\%$; see Table 1 for statistics per age window). An unconditional model found that 38% of the variance in partial face presence was between infants and the

Table 1 | Descriptive statistics (median and median absolute deviation [MAD]) for percentage of video frames containing partial and full faces across age windows

Age Window	Partial Face Occurrence			Full Face Occurrence		
	<i>n</i>	Median	MAD	<i>n</i>	Median	MAD
6–8 months	43	28%	16%	44	2%	2%
9–11 months	41	30%	16%	41	3%	2%
12–14 months	26	33%	11%	26	5%	3%
15–17 months	25	32%	8%	25	3%	3%

remaining 62% was within infants (i.e., between videos for the same infant). A conditional model—including the age window of each video as a linear predictor—found that each successive age window was associated with a 0.05 increase in the log-odds of a partial face (95% CI: [−0.03, 0.13]; OR = 1.05 [0.97, 1.14]), with 89.1% of the posterior mass above zero. The estimate points to a slight increase in partial-face presence across age, but the credible interval is consistent with no change (Fig. 1; Table S4, Supplementary information D). The second set of models revealed that the percentage of each video that contained a full face was much lower, ranging from 0.1% to 17% ($M = 4\%$, $SD = 3\%$; Table 1). An unconditional model found that 21% of the variance in full face presence was between infants, with the remaining 79% within infants. A conditional model found that full face presence increased modestly across age windows: $b = 0.11$ (95% CI: [0.01, 0.22]; OR = 1.12 [1.01, 1.25]), with 98.2% of the posterior mass above zero. The estimate points to a modest increase in full-face presence across age (Fig. 1; see also Table S5 in Supplementary information D).

Taken together, these analyses showed that full faces were rare in infants' view, that infants differed from one another in their exposure to faces, and that infants were increasingly exposed to full faces between 6 and 17 months of age.

Hypothesized expressive prototypes

We next examined infants' exposure to hypothesized expressive prototypes per emotion category. As depicted in Fig. 2, exposure to hypothesized expressive prototypes was quite rare overall. The total duration of any hypothesized expressive prototype across the 6- and 9-month age windows was under half an hour (27.9 min), which is only 0.6% of the total video frames and 17.6% of the video frames where a full face was present and codable. In other words, hypothesized expressive prototypes did not occur during the vast majority (over 82%) of full-face exposures. Even the most frequent category of hypothesized expressive prototypes, HAPPINESS, occurred in fewer than 0.5% of the total video frames and around 12% of the codable full face video frames (i.e., less than 20 min of the 83.9 h of video from these time points).

Across the 12- and 15-month age windows, from which primary caregivers were largely excluded due to the presence of head-mounted cameras, exposure to hypothesized expressive prototypes was even rarer. The total duration of any hypothesized prototype was less than 6 min, 0.2% of total video frames or 13.0% of full-face frames. Hypothesized prototypes for HAPPINESS occurred in about 0.1% of total video frames and 7.6% of full-face frames, lasting less than 4 min altogether.

In both earlier and later age windows, what exposure there was to hypothesized expressive prototypes was concentrated in a few emotion categories. The six emotion categories with the highest exposure were HAPPINESS, CONTENTMENT, SADNESS, SURPRISE, FEAR, and DISGUST, and the nine emotion categories with no exposure were AMUSEMENT, AWE, COYNESS, EMBARRASSMENT, INTEREST, PAIN, PRIDE, SHAME, AND TRIUMPH.

Exploratory examination of context

To follow up on the above observations, we conducted an exploratory examination of the context in which a subset of prototypes for ANGER, DISGUST, FEAR, HAPPINESS, and SADNESS occurred (FACS Investigator's Guide³⁹).

Namely, we coded the 10-second window of video recording surrounding each occurrence of a prototype for the general context of the interaction (e.g., playing with toys together, baby crying during mealtime), whether this was positive or negative in valence, and whether this valence matched the normative valence for the associated emotion category (e.g., playing with toys is positive, as are hypothesized expressions of HAPPINESS, but not hypothesized expressions of FEAR). Using this broad coding scheme, we found that HAPPINESS was the only emotion category with hypothesized expressive prototypes that occurred in mostly matching contexts (see Fig. 3). Hypothesized prototypes for ANGER, DISGUST, FEAR, and SADNESS rarely occurred in contexts where negative emotions would be expected and, in fact, mostly occurred during playtime. Additional descriptive statistics for the coded contexts and valences are available in Supplementary information E.

Observed facial configurations

Given the infrequency of hypothesized expressive prototypes for specific emotion categories, we sought to characterize the facial input infants did receive in terms of raw facial configurations. Across the 6- and 9-month age windows, faces were often still, with only 61.5% of full-face time (1.9% session time) containing any coded AU activity. When faces were moving, infants were exposed to 891 distinct facial configurations. This input was strongly concentrated: the 11 configurations that each accounted for at least 1% of full-face time (Table 2) together comprised roughly half of all occurrences (49.0%), while the remaining 880 configurations were individually rare and formed a long tail. The most frequent configurations were dominated by smiles (AU 12 alone and combined with mouth opening AU 25, AU 26 and cheek raising AU 6) and brow raises (AU 1 + AU 2; for visualizations, see Fig. 4).

Configurations differed markedly in how broadly they were distributed across infants. Overall, 473 configurations (53.1%) occurred in only a single infant's environment, and the median configuration was seen in just one infant. In contrast, each of the 11 most frequent configurations was observed across a large majority of the 44 infants (range 36–44; Table 2) and was produced by many distinct individuals. Infants' facial input thus combined a small set of near-ubiquitous, high-frequency configurations with a large, highly idiosyncratic tail.

Finally, we examined how these observed configurations related to the hypothesized expressive prototypes. Overlap was common but neither complete nor exclusive. Observed configurations mapped onto hypothesized prototypes with moderate or strong correspondence during 61.3% of full-face time (1.9% session time; Table 3). However, facial movements often did not correspond perfectly with any emotion category or corresponded with more than one emotion category. Out of the 11 most common facial configurations, six (55%) had match scores below 1 and six (55%) matched to multiple categories—including categories differing on key affective dimensions such as valence (e.g., AU 1 + AU 2 matched 0.80 with FEAR, INTEREST, and SURPRISE). Further, observed facial configurations with moderate-strong correspondence were most often simple, comprised of only one or two AUs (Fig. 5; see also Supplementary information F). Elaborate configurations were near-absent, with configurations of 5+ AUs accounting for only 6.6% of moving face time.

Discussion

The facial configurations that infants encounter in typical interactions with caregivers and family members form a basis for their understanding of emotion. To fully understand how infants learn about emotion, the facial information they receive must be quantified and described. Here, we examined infants' natural exposure to facial configurations by leveraging an existing, longitudinal corpus of video-recordings of 6- to 17-month-old infants' interactions with close others. In our first analysis, we examined how often faces were available in infants' field of view. We found that infants were generally exposed to far more partial than full faces, and that exposure to full faces was quite low but increased over the time periods observed. In our second analysis, we examined whether full faces contained facial

Fig. 1 | Presence of partial and full faces in infants' field of view across age windows. Error bands are 95% credible intervals. Sample size by window: 6-8 months, $n = 44$; 9-11 months, $n = 41$; 12-14 months, $n = 26$; 15-17 months, $n = 25$.

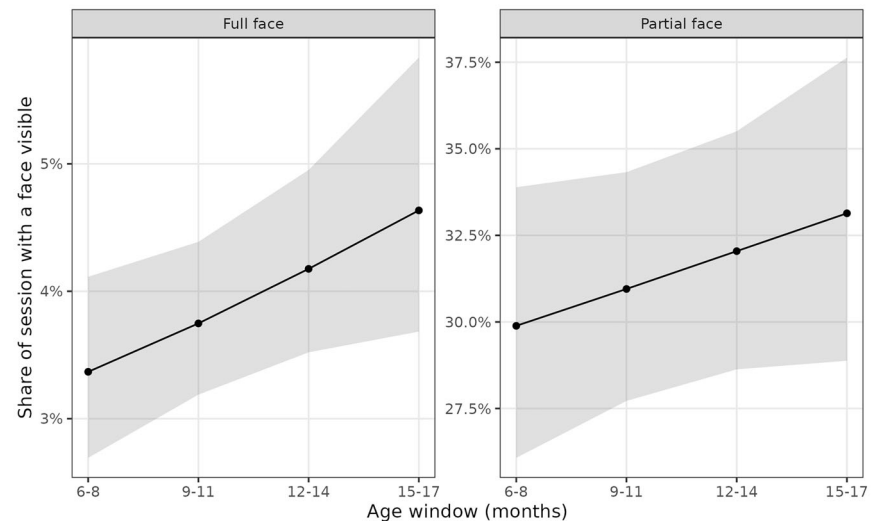
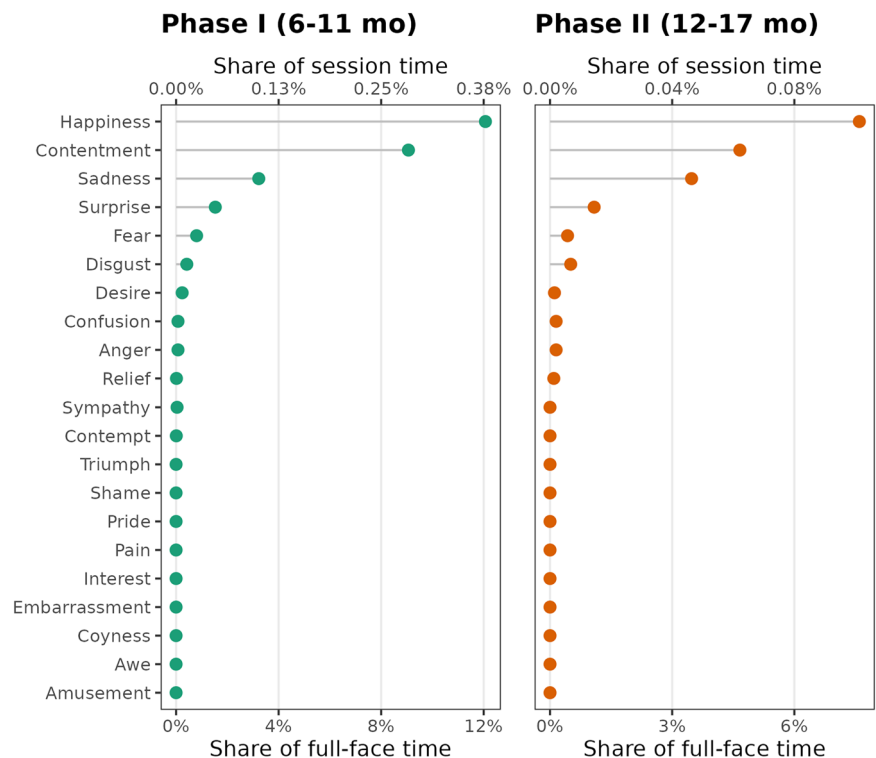


Fig. 2 | Infants' exposure to hypothesized expressive prototypes across age windows. The percentage of frames containing a facial configuration matching a hypothesized expressive prototype, for 6- and 9-month age windows (turquoise, left) and for 12- and 15-month age windows (orange, right). The double x-axis shows both the % of valid session frames (top) and the % of valid full-face frames (bottom).



movements matching any expression that has been hypothesized as a prototype for 21 emotion categories. We found that exposure was extremely rare—even HAPPINESS, the most-encountered category of hypothesized expressive prototype, occurred in fewer than 0.5% of all video frames. Moreover, when categories of hypothesized expressive prototypes were encountered, they often occurred in unexpected contexts, suggesting that they did not necessarily serve as a cue to the presumed emotional meaning. Descriptive analyses of the facial movement data further revealed that infants' facial input was dominated by a few, mostly simple facial configurations that did not correspond fully or uniquely to hypothesized category prototypes. These results build on evidence that emotions are expressed heterogeneously in daily life^{2,3,28}, such that infants must learn much more than one-to-one mappings between facial configurations and context/behavior to acquire expressive vocabularies for emotion.

Results suggest that infants at between 6 and 17 months are exposed to more partial faces (around 32% of the time, range 3% to 65%) than full faces (around 4% of the time, range 0.1% to 17%), which is broadly consistent with extant literature^{34,35}. Unlike prior work, however, we observed that exposure to full faces increased between 6 and 17 months of age; in those other studies, exposure to faces decreased from 1 to 3 months³⁵ or from 1 to 24 months³⁴. Several differences between our work and previous studies may account for this difference in findings. For example, it is possible that the change in face exposure varies with infant age, and our observed increase was a function of the developmental time points sampled. Alternatively, the difference in proportion of faces may reflect differences in the length or context of recording. Whereas Fausey and colleagues³⁴ collected video of up to 6 h (average 4.22) of unconstrained daily activities (i.e., the camera was embedded in a cap that the child wore), the parents and infants in

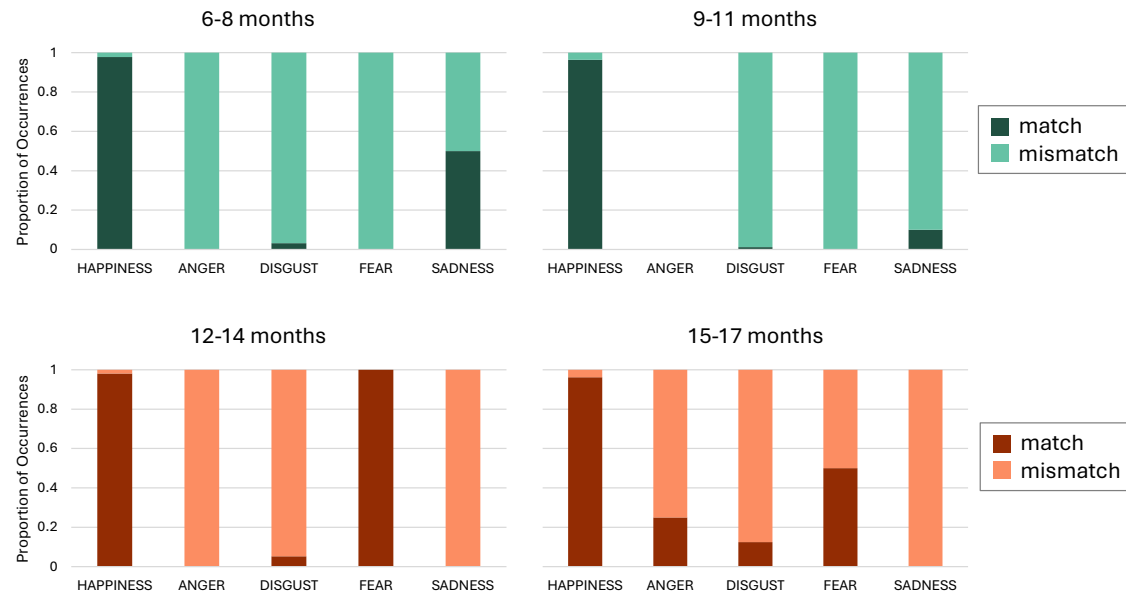


Fig. 3 | Proportion of occurrences of a subset of hypothesized expressive prototypes matching the valence of the surrounding context. Panels provide proportions for each age window, organized into 6- and 9-month windows (turquoise) versus 12- and 15-month windows (orange).

Table 2 | Observed configurations in infants' facial input in 6- and 9-month age windows

Configuration	AUs	Events	Infants	Individuals	Duration (s)	% Full-Face Time	% Session Time	Best Match Score	Hypothesized Prototype(s)
12	1	1236	44	93	861.16	9.06%	0.28%	1.00	CONTENTMENT, HAPPINESS
12 + 25 + 26	3	553	43	81	313.32	3.30%	0.10%	0.86	DESIRE, RELIEF
4	1	506	42	69	300.18	3.16%	0.10%	1.00	SADNESS
12 + 25	2	524	42	84	285.92	3.01%	0.09%	1.00	HAPPINESS
1 + 2	2	427	41	72	249.19	2.62%	0.08%	0.80	FEAR, INTEREST, SURPRISE
6 + 12	2	226	38	57	202.88	2.14%	0.07%	1.00	HAPPINESS
1 + 2 + 12	3	275	39	62	183.86	1.94%	0.06%	1.00	INTEREST
6 + 12 + 25 + 26	4	197	36	54	153.03	1.61%	0.05%	0.89	AMUSEMENT, COYNESS, HAPPINESS, TRIUMPH
17	1	274	38	65	116.76	1.23%	0.04%	0.67	DISGUST, SADNESS, SHAME
14	1	232	37	62	102.74	1.08%	0.03%	0.67	CONTEMPT
24	1	241	38	64	95.04	1.00%	0.03%	0.50	ANGER, SYMPATHY
Remaining 880 configurations (each <1%)	—	6386	—	—	2,979.71	31.36%	0.99%	—	

Configurations occupying $\geq 1\%$ of full-face time. Duration is person-time; co-visible faces contribute independently. *Individuals* is distinct infant-by-identity combinations. *Hypothesized Prototype(s)* lists every emotion category achieving the best match score (highest Dice coefficient); a configuration matching several reflects overlap among the hypothesized sets and these are not additive.

SEEDLingS were observed during 1-hour sessions with experimenters and additional recording equipment present. In Fausey's data³⁴, infant-adult (or -other) interactions may have been a relatively smaller proportion of the longer recording sessions from each child compared to the proportion of interactions in our shorter recording sessions. In addition, the study by Fausey³⁴ was cross-sectional, whereas the dataset for the present study was longitudinal; thus, with our modeling approach, we were able to demonstrate that exposure increased *in the same infants* across development. It could be that apparent differences in exposure trajectory are artifacts of the method, and Fausey's cross-sectional results reflect, at least in part, differences in parental attention to the recording devices at different ages, while parents in our longitudinal sample may have become more comfortable with the recording devices over time.

What can explain the observed variation in face exposure across infants, and change across infancy? The large between-participants range in face views is likely due to the activities each infant experienced. There were

no set instructions for SEEDLingS' naturalistic interactions, and so infants who were very active, or in bouncy chairs for instance, had a different perspective on their caregivers than infants who were less active or sitting on the floor. Additionally, the number of people interacting with infants varied by family and from visit to visit. Changes in face exposure over time may be driven in part by infants' developing motor abilities. During the unconstrained everyday activities that unfold over the 4-to-6-hour period recorded in prior work, face exposure may vary as infants develop increased motor control over their visual input^{33,34}. In our shorter sessions, the focus may have been more on parent-child interaction, and the face may become more central over time, as infants begin to participate more actively in social interactions.

Our primary finding is that infants are exposed to a wide variety of facial configurations that do not only include hypothesized expressive prototypes for emotions. Several results support this conclusion. First, the specific combinations of facial movements hypothesized to serve as



Fig. 4 | Visualization of observed facial configurations that accounted for at least 1% of full-face time in 6- and 9-month age windows. Configurations were visualized in FaceGen Modeler (version 3.34; Singular Inversions, Inc., Toronto, Ontario, Canada). All intensities were set to 100%, unless this yielded visible distortions (e.g., for AUs 1 and 2).

Table 3 | Distribution of match scores of observed facial configurations against hypothesized expressive prototypes

Match Score Band	Configurations	Events	Duration (s)	% Configurations (Unweighted)	% Full-Face Time	% Session Time
none (<0.2)	8	18	7.5	0.9%	0.1%	0.0%
weak (0.2–0.4)	3	15	7.5	0.3%	0.1%	0.0%
moderate (0.4–0.7)	285	2839	1247.9	32.0%	13.1%	0.4%
strong (≥0.7)	595	8205	4580.8	66.8%	48.2%	1.5%
no configuration coded	—	—	3657.2	—	38.5%	1.2%

Best-match Dice coefficient per observed configuration, organized by interpretative bands from ref. ⁵³. Session percentages do not sum to 100%; the remainder is time with no fully visible face.

expressive prototypes for many categories of emotion—including HAPPINESS, SADNESS, and SURPRISE—are relatively rare in infants’ visual environments. This finding is somewhat surprising given infants’ early emerging abilities to discriminate between and categorize such hypothesized expressive prototypes^{19,20}. Researchers typically use such configurations when assessing infants’ emotion perception, presumably based on an assumption that these hypothesized prototypes will be common. However, there is little evidence to support this assumption⁴, even in studies of adults. In the meta-analysis mentioned earlier, for example, adults were found to only scowl about 35% with ANGER elicitors and to smile about 20% of the time with HAPPINESS elicitors³. Other observational evidence suggests a lower prevalence of hypothesized prototypes: McDuff and colleagues⁵⁴ estimated the average base rate of smiles to be 5.62% and brow furrows just 2.84% in the spontaneous facial movement of adults watching online videos. The frequency of hypothesized expressive prototypes has yet to be assessed in adult-adult interactions—work that could be accomplished by leveraging recent conversational datasets⁵⁵.

Second, exploratory analyses revealed that hypothesized expressive prototypes for purportedly ‘basic’ emotions of ANGER, DISGUST, FEAR, and SADNESS routinely occurred in positively-valenced situations, such as play-time, rather than alongside more congruent behaviors. In one example, a mother made a wide-eyed gasping (‘scared’) face before telling the infant

excitedly that her brother was entering the room. Thus, even when infants were exposed to hypothesized expressive prototypes, they occurred in contexts that may not be consistent with theoretical expectations. Combined with the relative rarity of hypothesized expressive prototypes, this finding suggests that caregivers’ facial configurations may not be intended to communicate a particular emotion, but rather may serve a different, attention-grabbing function⁵⁶.

Third, our infants were frequently exposed to combinations of facial movements that did not overlap fully or uniquely with hypothesized expressive prototypes. The findings from our observed facial configuration analysis revealed that infants’ collective input was dominated by a few common movements and combinations thereof. The multi-movement configurations typically depicted in the emotion literature^{11,39,40,42} were very rare, and most observed configurations occurred for only a single infant. Stronger correspondence with hypothesized expressive prototypes was concentrated among the simplest configurations, such as smiling (AU 12) or brow raises (AU 1 + AU 2). Even though individual observed facial configurations typically had moderate to strong correspondence with hypothesized prototypes in general, these prototypes often spanned multiple emotion categories. Equally critically, we cannot know from the present data whether these facial configurations were true reflections of internal feelings

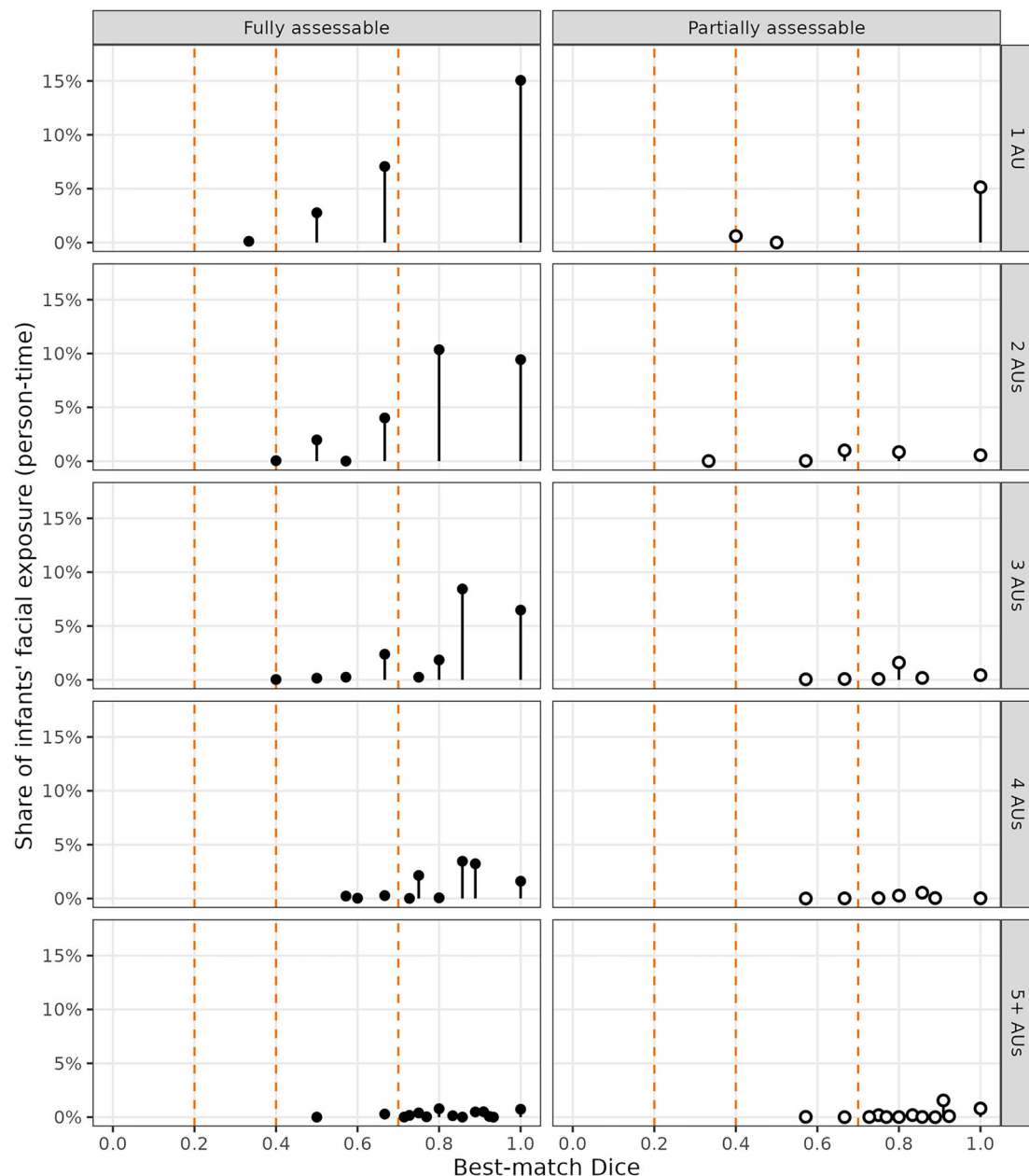


Fig. 5 | Match scores of observed facial configurations against all hypothesized expressive prototypes, broken down by match assessability and configuration size. Panels provide distributions of match scores between observed facial configurations and hypothesized expressive prototypes. Distributions are plotted separately for each configuration size (1–5+ AUs; rows) and assessability

level (fully vs partially assessable; columns). Heights are shares of infants' total facial exposure across all ten panels, such that columns do not each sum to 100%. Vertical bands (orange dash) reflect thresholds for no (≤ 0.2), weak (0.2–0.4), moderate (0.4–0.7), and strong correspondence (≥ 0.7).

or rather as communicative signals in the absence of the hypothesized emotion⁵⁷.

Taken together, the fact that hypothesized expressive prototypes for specific emotion categories were relatively infrequent and that infants instead saw an idiosyncratic assortment of simple configurations that overlapped with hypothesized prototypes from multiple categories suggests that the distribution of facial input infants receive is highly variable and may not be directly related to emotion. Infants do form prototypes based on this distribution of input, but those prototypes are probabilistic and idiographic, based on personal experience, rather than universal⁵⁸. The combinations of facial movements that we identified in the SEEDLingS corpus did bear a family resemblance to each other⁵⁹ and so may form a basis for these infants' social and emotional learning. This is not to say that infants are using facial

configurations solely for the purpose of inferring others' emotions. Facial movements are not inherently 'emotional' and often do not have an emotional meaning (e.g., when a furrowed brow comes from looking toward the sun)^{57,60}. What infants are doing, first and foremost, is learning to segment and integrate across facial movements for the purpose of predicting what comes next in their social and physical environment—regardless of these predictions' correspondence to categories for emotion⁴. Future work can probe the complexity of infants' developing models of facial behavior by examining the dynamics of specific facial movements (whether defined using action units^{41,61} or other types of landmarks/features^{62,63}). With that said, the low overall exposure to full faces (the basis for assessing hypothesized expressive prototype occurrence and describing observed facial configurations) suggests that infants are not getting a lot of emotion-related

information from the face in general. Future work is needed to analyze these kinds of naturalistic interactions multimodally, to assess the contribution of voice, bodily movement, etc. to infants' emotional learning, and the extent to which each makes unique or redundant contributions.

Alongside an exploration of multimodal patterns in infants' emotional input, the specific role of language should also be assessed. Theoretical advances suggest that emotion labels play a critical role in the development of emotion understanding, enabling infants to learn which signals serve as information for which emotion categories^{30,64–66}. These proposals have received support from recent experimental work^{67–69}, yet little is known about how often parents and caregivers use emotion labels in the presence of their infants, and what kinds of words these are (for initial evidence, see ref. 70). Using data sets such as SEEDLingS, it would be possible to examine the relative frequency of emotion labels in naturalistic interactions across a key period in infant development. Critically, as the strongest opportunities for learning emotion categories may come from instances where facial configurations and emotion labels are presented concordantly⁷¹, it would also be possible to assess the probabilistic intersection between these two channels of input. These and other analyses await future work.

Limitations

The present findings should be interpreted in light of several limitations. As mentioned above, the specific context of the interactions (i.e., recording sessions) and the age period sampled inherently temper the conclusions we can draw about infants' emotional facial input more broadly and how our observations would extend to different periods in development. Future work is needed to assess longer trends⁷², as well as the ecologies of other cultural and communicative environments⁷³. Additionally, caregiver-worn cameras occluded a substantial portion of full faces at later time points, such that these age windows could not be meaningfully included in longitudinal analyses surrounding exposure to hypothesized expressive prototypes or observed facial configurations. These analyses would be a particularly valuable contribution of future research; for example, it is possible that caregivers initially use more exaggerated, stereotypical facial configurations (as in 'motionese'²⁷ or 'motherese'²⁵) and that more variable facial configurations emerge as infants' communication skills grow more sophisticated and the interactional context grows richer. Finally, our behavioral coding did not include all facial and bodily movements that have been hypothesized to be relevant to emotion (e.g., hand movements).

Conclusion

Despite these considerations, the present study takes a substantive, longitudinal, and ecological approach to understanding the types of emotional input infants receive during naturalistic interactions and therefore has methodological and theoretical implications for our knowledge of emotional development in infancy. By investigating how hypothesized expressive prototypes stack up against real life, we have shown that the combinations of facial movements assumed to be at the core of emotion are neither common nor fixed to certain types of behavioral contexts. Instead, we find other facial vocabularies are used to communicate with infants. If this is the input infants most commonly receive, how are infants acquiring emotion categories? This work suggests next steps for developmental and emotion psychology that have the potential to fundamentally reframe our understanding of infants' learning problem.

Data availability

SEEDLingS data are available via Databrary. Further information about the data set, including access instructions, is available here: <https://doi.org/10.21415/T5PK6D>.

Code availability

Analytic code and supporting materials are available via a repository hosted by the Center for Open Science (OSF) at <https://osf.io/jmckl/>.

Received: 23 July 2025; Accepted: 9 September 2026;

Published online: 24 September 2026

References

1. Ruba, A. L. & Pollak, S. D. The development of emotion reasoning in infancy and early childhood. *Annu. Rev. Dev. Psychol.* **2**, 503–531 (2020).
2. Camras, L. A., Castro, V. L., Halberstadt, A. G. & Shuster, M. M. Facial expressions in children are rarely prototypical. in *The Psychology of Facial Expressions* (eds Russell, J. A. & Fernandez-Dols, J. M.) 279–296 (Cambridge University Press, New York, NY, 2017).
3. Durán, J. I. & Fernández-Dols, J.-M. Do emotions result in their predicted facial expressions? A meta-analysis of studies on the co-occurrence of expression and emotion. *Emotion* **21**, 1550–1569 (2021).
4. Barrett, L. F., Adolphs, R., Marsella, S., Martinez, A. M. & Pollak, S. D. Emotional expressions reconsidered: challenges to inferring emotion from human facial movements. *Psychol. Sci. Public Interest* **20**, 1–68 (2019).
5. Barrett, L. F. Context reconsidered: complex signal ensembles, relational meaning, and population thinking in psychological science. *Am. Psychol.* **77**, 894–920 (2022).
6. Camras, L. A. et al. Do infants show distinct negative facial expressions for fear and anger? Emotional expression in 11-month-old European American, Chinese, and Japanese infants. *Infancy* **11**, 131–155 (2007).
7. Camras, L. A. & Shutter, J. M. Emotional facial expressions in infancy. *Emot. Rev.* **2**, 120–129 (2010).
8. Le Mau, T. et al. Professional actors demonstrate variability, not stereotypical expressions, when portraying emotional states in photographs. *Nat. Commun.* **12**, 5037 (2021).
9. Ekman, P. Facial expressions. in *The Science of Facial Expression* (eds Fernandez-Dols, J. M. & Russell, J. A.) (Oxford University Press, 2017).
10. Izard, C. E. Basic emotions, natural kinds, emotion schemas, and a new paradigm. *Perspect. Psychol. Sci.* **2**, 260–280 (2007).
11. Keltner, D., Sauter, D., Tracy, J. L. & Cowen, A. S. Emotional expression: advances in basic emotion theory. *J. Nonverbal Behav.* <https://doi.org/10.1007/s10919-019-00293-3> (2019).
12. Barrett, L. F. & Theriault, J. What's real? A philosophy of science for social psychology. *Handbook of Social Psychology* (2025).
13. Farroni, T., Menon, E., Rigato, S. & Johnson, M. H. The perception of facial expressions in newborns. *Eur. J. Dev. Psychol.* **4**, 2–13 (2007).
14. Field, T. M., Woodson, R., Greenberg, R. & Cohen, D. Facial expression by neonates. *Annu. Prog. Child Psychiatry Child Dev.* **16**, 119–125 (1983).
15. Kotsoni, E., de Haan, M. & Johnson, M. H. Categorical perception of facial expressions by 7-month-old infants. *Perception* **30**, 1115–1125 (2001).
16. Mumme, D. L., Fernald, A. & Herrera, C. Infants' responses to facial and vocal emotional signals in a social referencing paradigm. *Child Dev.* **67**, 3219–3237 (1996).
17. Mumme, D. L. & Fernald, A. The infant as onlooker: learning from emotional reactions observed in a television scenario. *Child Dev.* **74**, 221–237 (2003).
18. Nelson, C. A., Morse, P. A. & Leavitt, L. A. Recognition of facial expressions by seven-month-old infants. *Child Dev.* 1239–1242 (1979).
19. Schwartz, G. M., Izard, C. E. & Ansel, S. E. The 5-month-old's ability to discriminate facial expressions of emotion. *Infant Behav. Dev.* **8**, 65–77 (1985).
20. Serrano, J. M., Iglesias, J. & Loeches, A. Visual discrimination and recognition of facial expressions of anger, fear, and surprise in 4-to 6-month-old infants. *Dev. Psychobiol. J. Int. Soc. Dev. Psychobiol.* **25**, 411–425 (1992).

21. Sorce, J. F., Emde, R. N., Campos, J. J. & Klinnert, M. D. Maternal emotional signaling: Its effect on the visual cliff behavior of 1-year-olds. *Dev. Psychol.* **21**, 195 (1985).
22. Widen, S. C. Children's interpretation of facial expressions: the long path from valence-based to specific discrete categories. *Emot. Rev.* **5**, 72–77 (2013).
23. Widen, S. C. & Russell, J. A. Children acquire emotion categories gradually. *Cogn. Dev.* **23**, 291–312 (2008).
24. Young-Browne, G., Rosenfeld, H. M. & Horowitz, F. D. Infant discrimination of facial expressions. *Child Dev.* 555–562 (1977).
25. Fernald, A. Four-month-old infants prefer to listen to motherese. *Infant Behav. Dev.* **8**, 181–195 (1985).
26. *Talking to Children: Language Input and Acquisition* (Cambridge University Press, Cambridge, 1977).
27. Brand, R. J., Baldwin, D. A. & Ashburn, L. A. Evidence for 'motionese': modifications in mothers' infant-directed action. *Dev. Sci.* **5**, 72–83 (2002).
28. Tcherkassof, A. & Dupré, D. The emotion–facial expression link: evidence from human and automatic expression recognition. *Psychol. Res.* <https://doi.org/10/ghmbwn> doi:10/ghmbwn. (2020)
29. LoBue, V. & Adolph, K. E. Fear in infancy: lessons from snakes, spiders, heights, and strangers. *Dev. Psychol.* **55**, 1889–1907 (2019).
30. Hoemann, K., Xu, F. & Barrett, L. F. Emotion words, emotion concepts, and emotional development in children: a constructionist hypothesis. *Dev. Psychol.* **55**, 1830–1849 (2019).
31. Hoemann, K. et al. Developing an understanding of emotion categories: Lessons from objects. *Trends Cogn. Sci.* **24**, 39–51 (2020).
32. Oakes, L. M. Plasticity may change inputs as well as processes, structures, and responses. *Cogn. Dev.* **42**, 4–14 (2017).
33. Smith, L. B., Jayaraman, S., Clerkin, E. & Yu, C. The developing infant creates a curriculum for statistical learning. *Trends Cogn. Sci.* **22**, 325–336 (2018).
34. Fausey, C. M., Jayaraman, S. & Smith, L. B. From faces to hands: changing visual input in the first two years. *Cognition* **152**, 101–107 (2016).
35. Sugden, N. A., Mohamed-Ali, M. I. & Moulson, M. C. I spy with my little eye: typical, daily exposure to faces documented from a first-person infant perspective. *Dev. Psychobiol.* **56**, 249–261 (2014).
36. Bergelson, E. Bergelson SEEDLingS HomeBank Corpus. <https://doi.org/10.21415/T5PK6D> (2017).
37. Sugden, N. A. & Moulson, M. C. Hey baby, what's "up"? One-and 3-month-olds experience faces primarily upright but non-upright faces offer the best views. *Q. J. Exp. Psychol.* **70**, 959–969 (2017).
38. Ekman, P. E., Friesen, W. V. & Hager, J. C. *Facial Action Coding System: A Technique for the Measurement of Facial Movement* (Research Nexus, 2002).
39. Ekman, P., Friesen, W. V. & Hager, J. C. *Facial Action Coding System Investigator's Guide* (Research Nexus, 2002).
40. Cordaro, D. T. et al. Universals and cultural variations in 22 emotional expressions across five cultures. *Emotion* **18**, 75–93 (2018).
41. Du, S., Tao, Y. & Martinez, A. M. Compound facial expressions of emotion. *Proc. Natl. Acad. Sci. USA* **111**, E1454–E1462 (2014).
42. Matsumoto, D., Keltner, D., Shiota, M. N., O'Sullivan, M. & Frank, M. Facial expressions of emotion. in *Handbook of Emotions* (eds Lewis, M., Haviland-Jones, J. M. & Barrett, L. F.) 211–234 (Guildford Press, New York, NY, 2008).
43. Laing, C. & Bergelson, E. Analyzing the effect of sibling number on input and output in the first 18 months. *Infancy* **29**, 175–195 (2024).
44. Kalenkovich, E. et al. A year of nouns from English-learning infants' daily lives: the SEEDLingS-Nouns dataset. *Behav. Res* **57**, 298 (2025).
45. Datavyu Team. *Datavyu: A Video Coding Tool* (Databrary Project, New York University, 2014).
46. Gelman, A. et al. *Bayesian Data Analysis* (CRC Press, 2014).
47. Flores, R. D. et al. Before/after Bayes: a comparison of frequentist and Bayesian mixed-effects models in applied psychological research. *Br. J. Psychol.* **113**, 1164–1194 (2022).
48. Figueroa-Zúñiga, J. I., Arellano-Valle, R. B. & Ferrari, S. L. Mixed beta regression: a Bayesian perspective. *Comput. Stat. Data Anal.* **61**, 137–147 (2013).
49. Bürkner, Paul-Christian. *Handle missing values with brms*
50. Barr, D. J., Levy, R., Scheepers, C. & Tily, H. J. Random effects structure for confirmatory hypothesis testing: keep it maximal. *J. Mem. Lang.* **68**, 255–278 (2013).
51. Bürkner, P.-C. & Charpentier, E. Modelling monotonic effects of ordinal predictors in Bayesian regression models. *Br. J. Math. Stat. Psychol.* **73**, 420–451 (2020).
52. Bürkner, P.-C. brms: an R package for Bayesian multilevel models using STAN. *J. Stat. Softw.* **80**, 1–28 (2017).
53. Haidt, J. & Keltner, D. Culture and facial expression: open-ended methods find more expressions and a gradient of recognition. *Cogn. Emot.* **13**, 225–266 (1999).
54. McDuff, D. et al. Large-scale observational evidence of cross-cultural differences in facial behavior. *J. Nonverbal Behav.* **41**, 1–19 (2017).
55. Reece, A. et al. The CANDOR corpus: insights from a large multimodal dataset of naturalistic conversation. *Sci. Adv.* **9**, eadf3197 (2023).
56. Trujillo, J. P. & Holler, J. Conversational facial signals combine into compositional meanings that change the interpretation of speaker intentions. *Sci. Rep.* **14**, 2286 (2024).
57. Crivelli, C. & Fridlund, A. J. Facial displays are tools for social influence. *Trends Cogn. Sci.* **22**, 388–399 (2018).
58. Rosch, E. Principles of categorization. in *Cognition and Categorization* (eds Rosch, E. & Lloyd, B. B.) 2–48 (Erlbaum, Hillsdale, NJ, 1978).
59. Posner, M. I. & Keele, S. W. On the genesis of abstract ideas. *J. Exp. Psychol.* **77**, 353–363 (1968).
60. Wnuk, E. & Wodowski, J. Culture shapes how we describe facial expressions. *Sci. Rep.* **14**, 21589 (2024).
61. Yan, Y. et al. The brain computes dynamic facial movements for emotion categorization using a third pathway. *Proc. Natl. Acad. Sci. USA* **122**, e2423560122 (2025).
62. Aleksic, P. S. & Katsaggelos, A. K. Automatic facial expression recognition using facial animation parameters and multistream HMMs. *IEEE Trans. Inf. Forensics Security* **1**, 3–11 (2006).
63. Wan, M. et al. InfAnFace: Bridging the infant–adult domain gap in facial landmark estimation in the wild. in *2022 26th International Conference on Pattern Recognition (ICPR)* 4486–4492 <https://doi.org/10.1109/ICPR56361.2022.9956647> (2022).
64. Barrett, L. F. *How Emotions Are Made: The Secret Life of the Brain* (Houghton Mifflin Harcourt, New York, NY, 2017).
65. Ruba, A. L., Pollak, S. D. & Saffran, J. R. Acquiring complex communicative systems: Statistical learning of language and emotion. *Top. Cogn. Sci.* **14**, 432–450 (2022).
66. Shablack, H. & Lindquist, K. A. The role of language in emotional development. in *Handbook of Emotional Development* 451–478 (Springer, 2019).
67. Ogren, M. & Sandhofer, C. M. Emotion words link faces to emotional scenarios in early childhood. *Emotion* **22**, 167–178 (2022).
68. Price, G. F., Ogren, M. & Sandhofer, C. M. Sorting out emotions: how labels influence emotion categorization. *Dev. Psychol.* **58**, 1665–1675 (2022).
69. Ruba, A. L., Meltzoff, A. N. & Repacholi, B. M. Superordinate categorization of negative facial expressions in infancy: the influence of labels. *Dev. Psychol.* <https://doi.org/10.1037/dev0000892> (2020).
70. Ogren, M. & Sandhofer, C. M. Emotion words in early childhood: a language transcript analysis. *Cogn. Dev.* **60**, 101122 (2021).
71. Gogate, L. J. Learning of syllable–object relations by preverbal infants: the role of temporal synchrony and syllable distinctiveness. *J. Exp. Child Psychol.* **105**, 178–197 (2010).

72. Widen, S. C. The development of children's concepts of emotion. in *Handbook of Emotions* (eds Barrett, L. F., Lewis, M. & Haviland-Jones, J. M.) 307–318 (Guildford Publications, 2016).
73. Keller, H. & Otto, H. The cultural socialization of emotion regulation during infancy. *J. Cross Cult. Psychol.* **40**, 996–1011 (2009).

Acknowledgements

We'd like to thank Dr. Erika Bergelson for sharing the SEEDLingS dataset with our research group.

Author contributions

M.O., L.L., and V.L. obtained, curated, and coded the SEEDLingS data. J.M.G., K.H., and L.L. analyzed the data and visualized the results. K.H. drafted the manuscript. K.H., J.M.G., M.O., L.L., L.F.B., L.O., and V.L. reviewed and revised the manuscript.

Funding

V.L. was supported by the National Science Foundation Grant #2333886 and the Eunice Kennedy Shriver National Institute of Child Health and Human Development Grant 1R21HD112313-01A1. L.F.B. was supported by the Unlikely Collaborators Foundation. The funders had no role in study design, data collection and analysis, decision to publish or preparation of the manuscript.

Competing interests

The authors declare no competing interests.

Additional information

Supplementary information The online version contains supplementary material available at <https://doi.org/10.1038/s44271-026-00539-z>.

Correspondence and requests for materials should be addressed to Katie Hoemann.

Peer review information *Communications Psychology* thanks Aina Puce and the other, anonymous, reviewer(s) for their contribution to the peer review of this work. Primary Handling Editors: Jennifer Bellinger. A peer review file is available.

Reprints and permissions information is available at <http://www.nature.com/reprints>

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Open Access This article is licensed under a Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License, which permits any non-commercial use, sharing, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if you modified the licensed material. You do not have permission under this licence to share adapted material derived from this article or parts of it. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by-nc-nd/4.0/>.

© The Author(s) 2026

This Article in Press is shared early to give you faster access to new research. It is citable and carries a permanent DOI. The final edited version will replace it automatically.