The representation of facial emotion expands from sensory to prefrontal cortex with development

Xiaoxu Fan[1], Abhishek Tripathi[2], Kelly R. Bijanki[1]

1. Baylor College of Medicine, Houston, TX, USA
2. Rice University, Houston, TX, USA

Corresponding author: Kelly R. Bijanki, bijanki@bcm.edu

Abstract

Facial expression recognition develops rapidly during infancy and improves from childhood to adulthood. As a critical component of social communication, this skill enables individuals to interpret others' emotions and intentions. However, the brain mechanisms driving the development of this skill remain largely unclear due to the difficulty of obtaining data with both high spatial and temporal resolution from young children. By analyzing intracranial EEG data collected from childhood (5-10 years old) and post-childhood groups (13-55 years old), we find differential involvement of high-level brain areas in processing facial expression information. For the post-childhood group, both the posterior superior temporal cortex (pSTC) and the dorsolateral prefrontal cortex (DLPFC) encode facial emotion features from a high-dimensional, continuous space. However, in children, the facial expression information is only significantly represented in the pSTC, not in the DLPFC. Further, the encoding of complex emotions in pSTC is shown to increase with age. Taken together, these data suggest that young children rely more on low-level sensory areas than on the prefrontal cortex for facial emotion processing, leading us to hypothesize that top-down modulation from prefrontal cortex to pSTC gradually matures during development to enable a full understanding of facial emotions, especially complex emotions which need social and life experience to comprehend.

Introduction

Understanding others' emotional states through their facial expressions is an important aspect of effective social interactions throughout the lifespan. Behavioral data suggest that facial emotion processing emerges very early in life[1,2], as infants just months old can distinguish happy and sad faces from surprised faces[3–5]. However, children's emotion recognition is substantially less accurate than adults, and this ability prominently improves across childhood and adolescence[6–11]. Although extensive research in cognitive and affective neuroscience has assessed developmental changes using behavioral and non-invasive neuroimaging approaches , our understanding of brain development related to facial expression perception remains limited.

One influential perspective on the development of face recognition is that it depends on the maturation of face-selective brain regions, including the fusiform face area (FFA), occipital face area (OFA), and posterior superior temporal sulcus (pSTS)[12]. Supporting this view, Gomez, et al. found evidence for

microstructural proliferation in the fusiform gyrus during childhood, suggesting that improvements in face recognition are a product of an interplay between structural and functional changes in the cortex[13]. Additionally, monkeys raised without exposure to faces fail to develop normal face-selective patches, suggesting that face experience is necessary for the development of the face-processing network[14]. It is likely that the gradual maturation of pSTS and FFA, two early sensory areas involved in the processing of facial expressions[15,16], contributes to the improved facial expression recognition over development. Yet, few studies have investigated the development of neural representation of emotional facial expressions in FFA and pSTS from early childhood to adulthood in human.

Besides the visual processing of facial configurations, understanding the emotional meaning of faces requires the awareness and interpretation of the emotional state of the other person, which is significantly shaped by life experience[17,18]. Thus, some researchers have proposed that the maturation of emotional information processing is related to the progressive increase in functional activity in the prefrontal cortex [6,19,20]. With development, greater engagement of the prefrontal cortex may facilitate top-down modulation of activity in more primitive subcortical and limbic regions, such as the amygdala[21–23]. Despite these theoretical advances, the functional changes in the prefrontal cortex during the perceptual processing of emotional facial expressions over development remains largely unknown.

Here, we analyze intracranial EEG (iEEG) data collected from childhood (5-10 years old) and post-childhood groups (13-55 years-old) while participants were watching a short audiovisual film. In our results, children's dorsolateral prefrontal cortex (DLPFC) shows minimal involvement in processing facial expression, unlike the post-childhood group. In contrast, for both children and post-childhood individuals, facial expression information is encoded in the pSTC, a brain region that contributes to the perceptual processing of facial expressions. Furthermore, the encoding of complex emotions in the pSTC increases with age. These neuroimaging data imply that social and emotional experiences shape the prefrontal cortex's involvement in processing the emotional meaning of faces throughout development, probably through top-down modulation of early sensory areas.

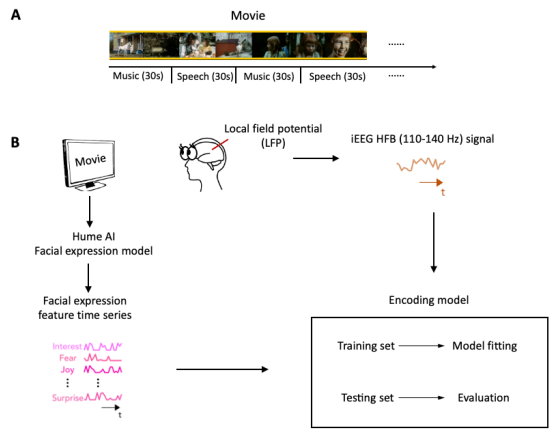

Figure 1. **Task design and analysis methods** . (**A**) Movie structure. A 6.5-minute short film was created by editing fragments from *Pippi on the Run* into a coherent narrative. The movie consisted of 13 interleaved blocks of videos accompanied by speech or music. (**B**) Data analysis schematic. Standard analysis pipeline for extracting emotion features from the movie and constructing encoding model to predict iEEG responses while participants watching the short film.

Results

Using AI and encoding models to study the neural representation of facial expression

In this study, we analyzed intracranial EEG (iEEG) data collected from a large group of human neurosurgical patients while they watched a short audiovisual film at the University Medical Center Utrecht[24]. The movie consisted of 13 interleaved blocks of videos accompanied by speech or music,

30 seconds each (**Figure 1A**). To characterize the neural representation of facial expression in the prefrontal cortex and low-level sensory areas across development, we analyzed iEEG data from 9 children (5-10 years old) and 31 post-childhood individuals (13-55 years old) who have electrode coverage in DLPFC, pSTC or both. First, Hume AI facial expression models were used to continuously extract facial emotion features from the movie (**Figure 1B**). Then, we tested how well encoding models constructed from the 48 facial emotion features (e.g., fear, joy) predict cortical high-frequency band (HFB) activity (110-140 Hz) induced by the presented movie (**Figure 1B**). The model performance was quantified as the correlation between the predicted and actual HFB activities, which is also called prediction accuracy.

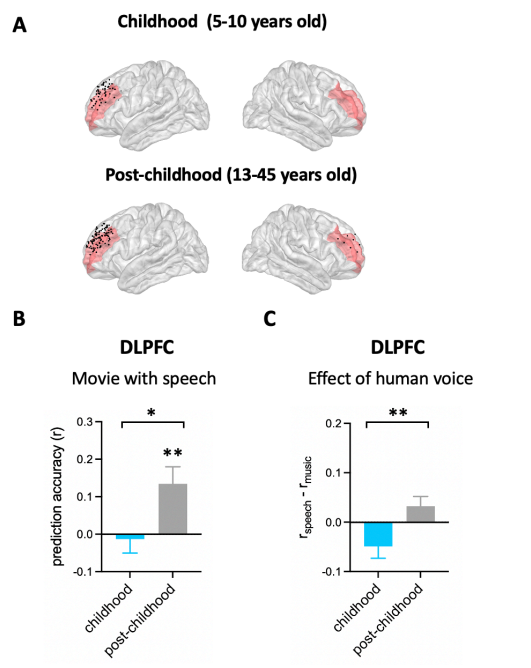

Figure 2. **Prediction performance of encoding models in DLPFC**. (**A**) Spatial distribution of electrodes in DLPFC. Electrodes in all participants from each group are projected onto MNI space and shown on the average brain. Red shaded areas indicate middle frontal cortex provided by the FreeSurfer Desikan-Killiany atlas[25]. Electrodes outside DLPFC are not shown. (**B**) The average prediction accuracy across participants for speech condition. The performance of encoding model is measured as Pearson correlation coefficient (r) between measured and predicted brain activities. (**C**) Prediction accuracy difference between speech condition and music condition for each group. Error bars are standard error of the mean. **P*<0.05; ***P*<0.01.

Differential representation of facial expression in children's DLPFC

Using the analysis approach described above, we examined how facial emotion information is represented by DLPFC (**Figure 2A**) while watching videos accompanied by speech (i.e. speech condition) in childhood and post-childhood groups. The prediction accuracy of the encoding model was significantly greater than zero in the post-childhood group (**Figure 2B**, *P*=0.0096, two-tailed permutation test), suggesting that the neural responses in DLPFC were dynamically modulated by the facial emotion features from the movie. However, facial emotion features were not encoded in children's DLPFC (**Figure 2B**, *P*=0.825, two-tailed permutation test). Moreover, the prediction accuracy in children's DLPFC was significantly lower than in the post-childhood group (*P*=0.0114, two-tailed permutation test). These findings show that the DLPFC dynamically encodes facial expression information in post-childhood individuals but not in young children.

To further understand the functional development of children's DLPFC, we compared the effect of human voice on the representation of facial expression in DLPFC between the two groups. The effect of human voice was quantified as difference in prediction accuracy between the speech and music conditions. Our results showed that human voice influences facial expression representation in the DLPFC differently across development (**Figure 2C**, *P*=0.0034, two-tailed permutation test). The presence

of human voice enhances facial expression representation in the DLPFC of post-childhood individuals but impairs it in children.

Taken together, there are significant developmental changes in DLPFC's involvement in facial expression perception.

The neural representation of facial expression in young children's pSTC

After identifying developmental differences in the involvement of high-level brain areas in processing facial expression, we next examined the neural representation of facial expression in children's early sensory areas. As an area in the core face network, posterior superior temporal sulcus (pSTS) has been associated with early stages of facial expression processing stream[12,15,26,27]. Although previous studies suggested that the development of facial recognition depends on the maturation of face-selective brain regions[13,14], it is still unclear how facial expression information is encoded in children's pSTS. Here, we examined the performance of the facial expression encoding model in a rare sample of two children (S19:8-year-old and S39: 5-year-old) with electrode coverage in pSTC (**Figure 3A**). In both cases, the encoding model significantly predicts the HFB neural signals in the pSTC under the speech condition (**Figure 3B**, $S19_{speech}$: P=0.0014, r=0.1951; $S39_{speech}$: P=0.0183, r=0.15). The prediction accuracy is reduced when human voice is absent from the video ($S19_{music}$: P=0.0313, r=0.1674; $S39_{music}$: P=0.3688, r=0.0574). Similarly, group-level results showed that the model performance is significantly greater than zero in the pSTC of post-childhood individuals (N=25, **Figure 3C** and **3D**, P=0.003, two-tailed permutation test) and this neural representation of facial expression information is significantly reduced when human voice is absent (paired-t-test, t_{24}=2.897,P=0.0079). These results provide evidence that children's sensory areas encode facial emotion features from a high-dimensional, continuous space in a manner similar to that of post-childhood individuals.

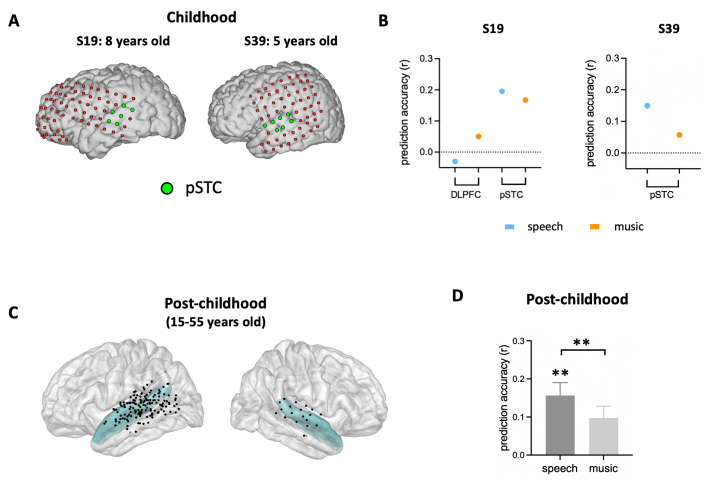

Figure 3. Prediction performance of encoding models in pSTC. (**A**) The electrode distribution of two children (s19 and s39). Electrodes in pSTC are green. (**B**) Prediction accuracy of encoding models in the two children. (**C**) Spatial distribution of recording contacts in post-childhood participants' pSTC. The pSTC electrodes identified in individual space are projected onto MNI space and shown on the average brain. Contacts other than pSTC are not shown. Blue shaded areas indicate superior temporal cortex provided by the FreeSurfer Desikan-Killiany atlas[25]. (**D**) Average prediction accuracy across post-childhood participants. Error bars are standard error of the mean. **P<0.01.

The complexity of facial expression encoding in the pSTC increases across development

To understand how facial expression representation in pSTC changes across development, we examined the feature weights of the facial expression encoding models in all participants with significant prediction

accuracy (10 post-childhood individuals and 2 children). The weight for each feature represents its relative contribution to predicting the neural response. First, we calculated the encoding weights for complex emotions (averaging guilt, embarrassment, pride, and envy, which were selected as the most representative complex emotions based on previous studies[28–30]) and basic emotions (averaging joy, sadness, fear, anger, disgust, and surprise). Then, we calculated their correlations with age separately. Our results showed that the encoding weight of complex emotion was significantly positively correlated with age (r_{12}=0.8512， *P*=0.004 , **Figure 4A** left). No significant correlation between encoding weight of basic emotion and age was observed (r_{12}=0.3913， *P*=0.2085, **Figure 4A** right). In addition, we computed

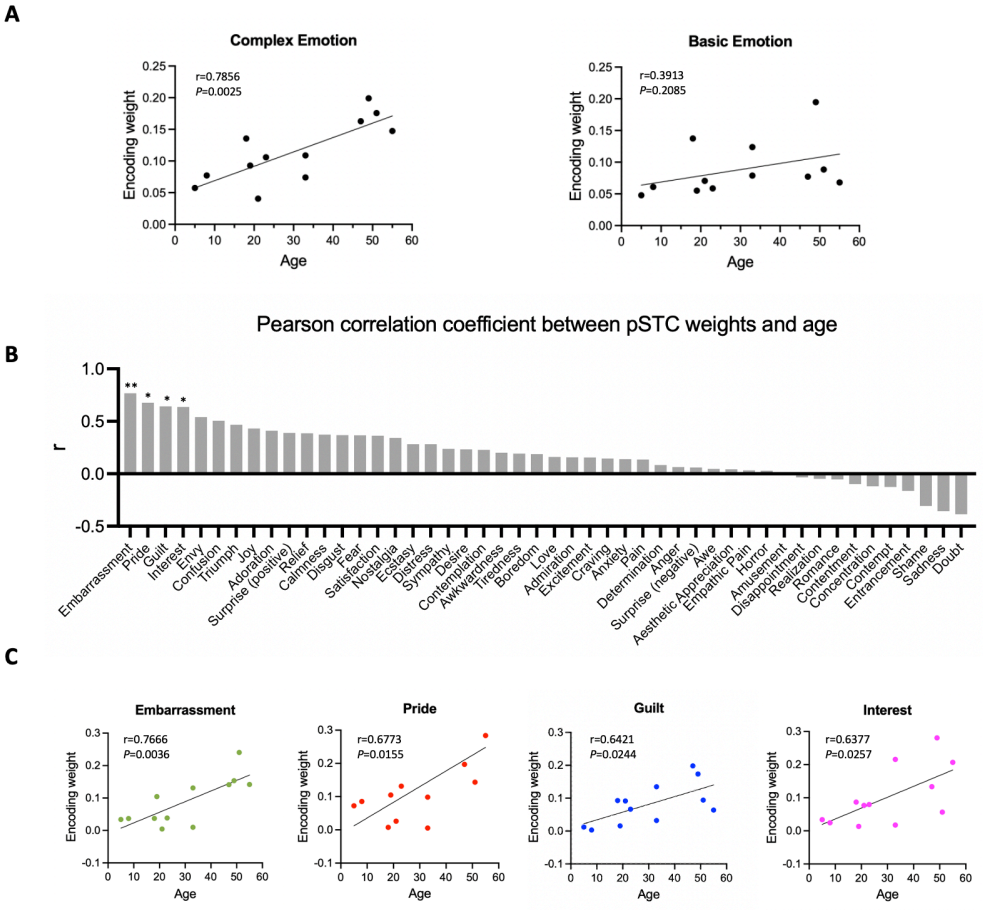

Figure 4. Correlation between encoding weights and age. (**A**) Left: Correlation between averaged encoding weights of five complex emotions and age. Right: Correlation between averaged encoding weights of six basic emotions and age. (**B**) Pearson correlation coefficient between encoding weights of 48 facial expression features and age. The results are ranked from largest to smallest. Significant correlations noted with *(*P*<0.05, uncorrected) or **(*P*<0.01, uncorrected). (**C**) Correlation between encoding weights of embarrassment, pride, guilt, interest and age (N=12).

Pearson correlations between each individual feature weight and age, ranking the r values from largest to smallest (**Figure 4B**). The highest correlations were found for embarrassment, guilt, pride, interest, and envy—emotions that are all considered complex emotion. Among them, the weights for embarrassment, guilt, pride, and interest showed significant positive correlations with age (**Figure 4C**, embarrassment: r=0.7666, *P*=0.0036; pride: r=0.6773, *P*=0.0155; guilt: r=0.6421, *P*=0.0244, interest:

r=0.6377, *P*=0.0257, uncorrected for multiple comparisons), suggesting that the encoding of these complex emotions in pSTC increases with age. Thus, our results suggest that as development progresses, the pSTC becomes increasingly engaged in encoding complex emotions which requires representing others' mental states and emerges later in development [31–33].

Methods

In this study, iEEG data from an open multimodal iEEG-fMRI dataset were analyzed[24].

Participants and electrode distribution

Due to the research purposes of the current study, only participants who had at least four electrode contacts in either DLPFC or pSTC were included in the data analysis (Table 1 and Table 2). Nine children (5-10 years old, 5 females) and thirty-one post-childhood individuals (13-55 years old, 18 females) are included in the present study. In the childhood group, eight participants had enough electrodes implanted in the DLPFC, and two had enough electrodes implanted in the pSTC. In the post-childhood group, thirteen participants had enough electrodes implanted in the DLPFC, and twenty-five had enough electrodes implanted in the pSTC.

Experimental procedures

A 6.5-minute short film was crafted by editing fragments from *Pippi on the Run* into a coherent narrative. The film is structured into 13 interleaved 30-second blocks of either speech or music, with seven blocks featuring background music only and six blocks retaining the original dialogue and voice from the video. Patients were asked to watch the movie while the intracranial EEG signals were recorded. No fixation cross was displayed in the middle of the screen or elsewhere. The movie was presented using the Presentation software (Neurobehavioral Systems, Berkeley, CA) and the sound was synchronized with the neural recordings. More data acquisition details can be found in Berezutskaya et al.'s article[24].

iEEG data processing

Electrode contacts and epochs contaminated with excessive artifacts and epileptiform activity were removed from data analysis by visual inspection. Raw data were filtered with a 50-Hz notch filter and re-referenced to the common average reference. For each electrode contact in each patient, the preprocessed data were band-pass filtered (110– 140 Hz, 4th-order Butterworth). The Hilbert transform was then applied to extract the analytic amplitude. Each event (block) was extracted in the 0 to 30 s time window around its onset. The fifth music block was excluded, as there were no faces presented on screen. Subsequently, the data were down-sampled to 400 Hz and square-root transformed. Finally, the data were normalized by z-scoring with respect to baseline periods (−0.2 to 0 s before stimulus onset).

Contact Location and Regions of Interest

We identified electrode contacts in STC in individual brains using individual anatomical landmarks (i.e., gyri and sulci). Superior temporal sulci and lateral sulci were used as boundaries. A coronal plane including the posterior tip of the hippocampus served as an anterior/posterior boundary. To identify electrode contacts in DLPFC, we projected the electrode contact positions provided by the open dataset

onto Montreal Neurological Institute-152 template brain (MNI) space, using FreeSurfer. DLPFC was defined based on the following sets of HCP-MMP1[34] labels on both left and right hemispheres: 9-46d, 46, a9-46v, and p9-46v.

Emotion feature extraction

Hume AI (https://www.hume.ai) was used to extract the facial emotion features from the video. When multiple faces appeared in the movie, the maximum score of the facial expression features across all faces was used for each emotion category. All the time courses of facial emotion features were resampled to 2Hz. No facial emotion features were extracted for the fifth music block due to the absence of faces. The full list of the 48 facial emotion features is shown in Figure 4B.

Encoding model fitting

To model iEEG responses to emotion, we used a linear regression approach with 48 facial emotion features extracted by Hume AI. Time-lagged versions of each feature (with 0, 0.5 and 1-second delays) were used in the model fitting. For each participant, high-frequency broadband (HFB) responses from all electrode contacts within each area were concatenated. To match the temporal resolution of the emotion feature time course, the HFB responses were binned into 500 ms windows. We then modeled the processed HFB response for each participant, each brain area, and each condition (speech vs music) using ridge regression. The optimal regularization parameter was assessed using 5-fold cross-validation, with the 20 different regularization parameters (log spaced between 10 and 10000). To keep the scale of the weights consistent, a single best overall value of the regularization coefficient was used for all areas in both the speech and music conditions in all patients. We used cross-validation iterator to fit the model and test it on held-out data. The model performance was evaluated by calculating Pearson correlation coefficients between measured and predicted HFB response of individual brain areas. The mean prediction accuracy (r value) of the encoding model with 5-fold cross-validation was then calculated. Non-parametric permutation tests were used to test whether the encoding model performance was significantly > 0 and whether there was a significant difference between groups. Specifically, we shuffled facial emotion feature data in time, and then we conducted the standard data analysis steps (described above) using the shuffled facial emotion features. This shuffle procedure was repeated 5000 times to generate a null distribution, and p-values were calculated as the proportion of results from shuffled data more extreme than the observed real value. A two-sided paired t-test was used to examine differences in encoding accuracy between speech and music conditions in post-childhood group (Figure 3D).

Weight analysis

To examine the correlation between encoding model weights and age, we obtained 48 encoding model weights from all folds of cross-validation for all participants whose pSTC significantly encoded facial expression (i.e. the p-value of prediction accuracy is less than 0.05). Thus, 10 post-childhood individuals and 2 children were involved in the weight analysis. The weight for each feature represents its relative contribution to predicting the neural response. A higher weight indicates that the corresponding feature has a stronger influence on neural activity, meaning that variations in this feature more significantly impact the predicted response. We used the absolute value of weights and therefore did not

discriminate whether facial emotion features were mapped to an increase or decrease in the HFB response.

Discussion

The current study examines functional changes in both low-level and high-level brain areas across development to provide valuable insights into the neural mechanisms underlying the maturation of facial expression perception. Based on our findings, we propose that young children rely primarily on early sensory areas rather than the prefrontal cortex for facial emotion processing. As development progresses, the prefrontal cortex becomes increasingly involved, perhaps serving to modulate responses in early sensory areas based on emotional context and enabling them to process complex emotions. This developmental progression ultimately enables the full comprehension of facial emotions in adulthood.

Behavioral results suggest that infants as young as only 7–8 months can categorize some emotions [3–5]. However, sensitivity to facial expressions in young children does not mean that they can understand the meaning of that affective state. For example, Kaneshige and Haryu (2014)[35] found that although 4-month-old infants could discriminate facial configurations of anger and happiness, they responded positively to both, suggesting that at this stage, they may lack knowledge of the affective meaning behind these expressions. This underscores the idea that additional processes need to be developed for children to fully grasp the emotional content conveyed by facial expressions. Although the neural mechanism behind this development is still unclear, a reasonable perspective is that it requires both visual processing of facial features and emotion-related processing for the awareness of the emotional state of the other person[17,36,37]. Indeed, growing evidence suggests that the prefrontal cortex plays an important role in integrating prior knowledge with incoming sensory information, allowing interpretation of the current situation in light of past emotional experience[19,38].

In the current study, we observed differential representation of facial expressions in the DLPFC between children and post-childhood individuals. First, in post-childhood individuals, neural activity in the DLPFC encodes high-dimensional facial expression information, whereas this encoding is absent in children. Second, while human voice enhances the representation of facial expressions in the DLPFC of post-childhood individuals, it instead reduces this representation in children. These results suggest that the DLPFC undergoes developmental changes in how it processes facial expressions. The absence of high-dimensional facial expression encoding in children implies that the DLPFC may not yet be fully engaged in emotional interpretation at an early age. Additionally, the opposite effects of human voice on facial expression representation indicate that multimodal integration of social cues develops over time. In post-childhood individuals, voices may enhance emotional processing by providing congruent information[39–41], whereas in children, the presence of voice might interfere with or redirect attentional resources away from facial expression processing [39,42,43].

There have been few neuroimaging studies directly examining the functional role of young children's DLPFC in facial emotion perception. Some evidence suggest that the prefrontal cortex continues to develop until adulthood to achieve its mature function in emotion perception[9,20,44], and for some emotion categories, this development may extend across the lifespan. For example, prefrontal cortex

activation during viewing fearful faces increases with age [18,19,45]. As there were not enough participants for us to calculate correlation between encoding model performance in DLPFC and age, it is still unclear whether the representation of facial expression in DLPFC increase linearly with age. One possibility is that the representation of facial expressions in the DLPFC gradually increases with age until it reaches an adult-like level. This would suggest a continuous developmental trajectory, where incremental improvements in neural processing accumulate over time. Another possibility is that development follows a more nonlinear pattern, showing improvement with prominent changes at specific ages. Interestingly, research has shown that performance on matching emotional expressions improves steadily over development, with notable gains in accuracy occurring between 9 and 10 years and again between 13 and 14 years, after which performance reaches adult-like levels [6].

Although there are only two children in our sample with enough electrodes in pSTC, our results clearly showed that facial expression is encoded in each child's pSTC. Moreover, the prediction accuracy of the encoding model in the two children was comparable to or higher than the average level in the post-childhood group. In the 5-year-old child (S19) who had electrode coverage in both DLPFC and pSTC, facial expressions were represented in the pSTC but not in the DLPFC. This rare and fortunate sampling allows us to rule out the possibility that the low prediction accuracy of the facial expression encoding model in the DLPFC is due to the reduced engagement in the movie-watching task for children. Consistent with our findings, previous studies have shown that the fusiform and superior temporal gyri are involved in emotion-specific processing in 10-year-old children[46]. Meanwhile, some other researchers found that responses to facial expression in the amygdala and posterior fusiform gyri decreased as people got older[20], but the use of frontal regions increased with age[44]. Therefore, we propose that early sensory areas like the fusiform and superior temporal gyri play a key role in facial expression processing in children, but their contribution may shift with age as frontal regions become more involved. Consistent with this perspective, our results revealed that the encoding weights for complex emotions in pSTC increased with age, suggesting a developmental trajectory in the neural representation of complex emotions in pSTC. This finding aligns with previous behavioral studies showing that social complex emotion recognition does not fully mature until young adulthood[31–33]. In fact, our results suggests that the representation of complex facial expressions in pSTC continues to develop over the lifespan. As for the correlation between basic emotion encoding and age, the lack of a significant effect in our study does not necessarily indicate an absence of developmental change but may instead be due to the limited sample size.

In summary, our study provides novel insights into the neural mechanisms underlying the development of facial expression processing. As with any study, several limitations should be acknowledged. First, most electrode coverage in our study was in the left hemisphere, potentially limiting our understanding of lateralization effects. Second, while our results provide insights into the role of DLPFC during development, we were unable to examine other prefrontal regions, such as the orbitofrontal cortex (OFC) and anterior cingulate cortex (ACC), to examine their unique contributions to emotion processing. Lastly, due to sample size constraints, we were unable to divide participants into more granular developmental stages, such as early childhood, adolescence, and adulthood, which could provide a more detailed characterization of the neural mechanisms underlying the development of facial expression

processing. Future studies using non-invasive methods, with more age-diverse samples will be essential for refining our understanding of how facial emotion processing develops across the lifespan.

Acknowledgement

We would like to thank Dr. Julia Berezutskaya for providing the audiovisual film that was used for iEEG data collection. This work was supported by fundings from United States National Institutes of Health (R01-MH127006).

Author contributions

X.F. and A.T. performed the data analysis. X.F. and K.R.B wrote the paper.

Competing interests

The author declares no competing interests.

Table 1. Demographic information of childhood group

ID	Sex	Age	Number of electrodes in left pSTC	Number of electrode in right pSTC	Number of electrode in left DLPFC	Number of electrode in right DLPFC
s02	F	9	0	0	9	0
s10	F	8	0	0	8	0
s19	F	8	6	0	8	0
s32	F	6	0	0	5	0
s37	M	5	0	0	4	0
s39	F	5	7	0	0	0
s41	M	7	0	0	4	0
s50	F	9	0	0	6	0
s63	M	5	0	0	9	0

Table 1. Demographic information of childhood group.

Table 2. Demographic information of post-childhood group

ID	Sex	Age	Number of electrodes in left pSTC	Number of electrode in right pSTC	Number of electrodes in left DLPFC	Number of electrode in right DLPFC
s1	M	55	0	4	0	0
s3	F	33	6	0	6	0
s5	F	33	8	0	8	0
s6	F	43	9	0	0	0
s12	M	37	0	0	12	0
s14	F	18	4	0	8	0
s16	M	17	8	0	9	0
s17	M	28	0	8	0	0
s18	F	15	6	0	11	0
s20	F	25	9	0	0	0
s22	M	21	4	0	0	0
s24	F	47	11	0	0	0
s25	M	14	0	0	6	0
s26	F	48	11	0	0	0
s27	M	15	11	0	0	0
s28	M	21	0	6	0	0
s31	F	13	0	0	7	0
s34	F	51	5	0	0	0
s38	F	14	0	0	4	0
s40	M	49	6	0	0	0
s43	M	19	0	4	0	8
s45	M	19	0	0	4	0
s48	F	18	5	0	0	0
s51	M	46	7	0	9	0
s54	F	31	4	0	0	0
s55	F	23	6	0	0	0
s57	F	36	0	0	7	0
s58	F	16	8	0	0	0
s59	F	30	7	0	0	0
s60	M	42	8	0	0	0
s61	F	16	8	0	0	0

Table 2. **Demographic information of post-childhood group.**

Reference

1. Barrera, M.E., and Maurer, D. (1981). The perception of facial expressions by the three-month-old. Child Dev. *52*, 203–206. 10.2307/1129231.

2. Walker-Andrews, A.S. (1997). Infants' perception of expressive behaviors: differentiation of multimodal information. Psychol. Bull. *121*, 437–456. 10.1037/0033-2909.121.3.437.

3. Caron, R.F., Caron, A.J., and Myers, R.S. (1982). Abstraction of invariant face expressions in infancy. Child Dev. *53*, 1008–1015.

4. Nelson, C.A., Morse, P.A., and Leavitt, L.A. (1979). Recognition of facial expressions by seven-month-old infants. Child Dev. *50*, 1239–1242. 10.2307/1129358.

5. Nelson, C.A., and Dolgin, K.G. (1985). The generalized discrimination of facial expressions by seven-month-old infants. Child Dev. *56*, 58–61.

6. Kolb, B., Wilson, B., and Taylor, L. (1992). Developmental changes in the recognition and comprehension of facial expression: implications for frontal lobe function. Brain Cogn. *20*, 74–84. 10.1016/0278-2626(92)90062-q.

7. Romani-Sponchiado, A., Maia, C.P., Torres, C.N., Tavares, I., and Arteche, A.X. (2022). Emotional face expressions recognition in childhood: developmental markers, age and sex effect. Cogn. Process. *23*, 467–477. 10.1007/s10339-022-01086-1.

8. Rodger, H., Vizioli, L., Ouyang, X., and Caldara, R. (2015). Mapping the development of facial expression recognition. Dev. Sci. *18*, 926–939. 10.1111/desc.12281.

9. Thomas, L.A., De Bellis, M.D., Graham, R., and LaBar, K.S. (2007). Development of emotional facial recognition in late childhood and adolescence. Dev. Sci. *10*, 547–558. 10.1111/j.1467-7687.2007.00614.x.

10. Johnston, P.J., Kaufman, J., Bajic, J., Sercombe, A., Michie, P.T., and Karayanidis, F. (2011). Facial emotion and identity processing development in 5- to 15-year-old children. Front. Psychol. *2*, 26. 10.3389/fpsyg.2011.00026.

11. Lawrence, K., Campbell, R., and Skuse, D. (2015). Age, gender, and puberty influence the development of facial emotion recognition. Front. Psychol. *6*, 761. 10.3389/fpsyg.2015.00761.

12. Duchaine, B., and Yovel, G. (2015). A revised neural framework for face processing. Annu. Rev. Vis. Sci. *1*, 393–416. 10.1146/annurev-vision-082114-035518.

13. Gomez, J., Barnett, M.A., Natu, V., Mezer, A., Palomero-Gallagher, N., Weiner, K.S., Amunts, K., Zilles, K., and Grill-Spector, K. (2017). Microstructural proliferation in human cortex is coupled with the development of face processing. Science *355*, 68–71. 10.1126/science.aag0311.

14. Arcaro, M.J., Schade, P.F., Vincent, J.L., Ponce, C.R., and Livingstone, M.S. (2017). Seeing faces is necessary for face-domain formation. Nat. Neurosci. *20*, 1404–1412. 10.1038/nn.4635.

15. Bernstein, M., and Yovel, G. (2015). Two neural pathways of face processing: A critical evaluation of current models. Neurosci. Biobehav. Rev. *55*, 536–546. 10.1016/j.neubiorev.2015.06.010.

16. Engell, A.D., and Haxby, J.V. (2007). Facial expression and gaze-direction in human superior temporal sulcus. Neuropsychologia *45*, 3234–3241. 10.1016/j.neuropsychologia.2007.06.022.

17. Pereira, M.R., Barbosa, F., de Haan, M., and Ferreira-Santos, F. (2019). Understanding the development of face and emotion processing under a predictive processing framework. Dev. Psychol. *55*, 1868–1881. 10.1037/dev0000706.

18. Yurgelun-Todd, D.A., and Killgore, W.D.S. (2006). Fear-related activity in the prefrontal cortex increases with age during adolescence: a preliminary fMRI study. Neurosci. Lett. *406*, 194–199. 10.1016/j.neulet.2006.07.046.

19. Williams, L.M., Brown, K.J., Palmer, D., Liddell, B.J., Kemp, A.H., Olivieri, G., Peduto, A., and Gordon, E. (2006). The mellow years?: neural basis of improving emotional stability over age. J. Neurosci. *26*, 6422–6430. 10.1523/JNEUROSCI.0022-06.2006.

20. Tessitore, A., Hariri, A.R., Fera, F., Smith, W.G., Das, S., Weinberger, D.R., and Mattay, V.S. (2005). Functional changes in the activity of brain regions underlying emotion processing in the elderly. Psychiatry Res. *139*, 9–18. 10.1016/j.pscychresns.2005.02.009.

21. Yurgelun-Todd, D. (2007). Emotional and cognitive changes during adolescence. Curr. Opin. Neurobiol. *17*, 251–257. 10.1016/j.conb.2007.03.009.

22. Hariri, A.R., Bookheimer, S.Y., and Mazziotta, J.C. (2000). Modulating emotional responses: effects of a neocortical network on the limbic system. Neuroreport *11*, 43–48. 10.1097/00001756-200001170-00009.

23. Hariri, A.R., Mattay, V.S., Tessitore, A., Fera, F., and Weinberger, D.R. (2003). Neocortical modulation of the amygdala response to fearful stimuli. Biol. Psychiatry *53*, 494–501. 10.1016/s0006-3223(02)01786-9.

24. Berezutskaya, J., Vansteensel, M.J., Aarnoutse, E.J., Freudenburg, Z.V., Piantoni, G., Branco, M.P., and Ramsey, N.F. (2022). Open multimodal iEEG-fMRI dataset from naturalistic stimulation with a short audiovisual film. Sci. Data *9*, 91. 10.1038/s41597-022-01173-0.

25. Desikan, R.S., Ségonne, F., Fischl, B., Quinn, B.T., Dickerson, B.C., Blacker, D., Buckner, R.L., Dale, A.M., Maguire, R.P., Hyman, B.T., et al. (2006). An automated labeling system for subdividing the human cerebral cortex on MRI scans into gyral based regions of interest. Neuroimage *31*, 968–980. 10.1016/j.neuroimage.2006.01.021.

26. Flack, T.R., Andrews, T.J., Hymers, M., Al-Mosaiwi, M., Marsden, S.P., Strachan, J.W.A., Trakulpipat, C., Wang, L., Wu, T., and Young, A.W. (2015). Responses in the right posterior superior temporal sulcus show a feature-based response to facial expression. Cortex *69*, 14–23. 10.1016/j.cortex.2015.03.002.

27. Fan, X., Wang, F., Shao, H., Zhang, P., and He, S. (2020). The bottom-up and top-down processing of faces in the human occipitotemporal cortex. eLife *9*. 10.7554/eLife.48764.

28. Alba-Ferrara, L., Hausmann, M., Mitchell, R.L., and Weis, S. (2011). The neural correlates of emotional prosody comprehension: disentangling simple from complex emotion. PLoS ONE *6*, e28701. 10.1371/journal.pone.0028701.

29. Burnett, S., Thompson, S., Bird, G., and Blakemore, S.-J. (2011). Pubertal development of the understanding of social emotions: Implications for education. Learn. Individ. Differ. *21*, 681–689. 10.1016/j.lindif.2010.05.007.

30. Russell, J.A., and Paris, F.A. (1994). Do Children acquire Concepts for Complex Emotions Abruptly? Int. J. Behav. Dev. *17*, 349–365. 10.1177/016502549401700207.

31. Garcia, N.V., and Scherf, K.S. (2015). Emerging sensitivity to socially complex expressions: A unique role for adolescence? Child Dev. Perspect. *9*, 84–90. 10.1111/cdep.12114.

32. Burnett, S., Bird, G., Moll, J., Frith, C., and Blakemore, S.-J. (2009). Development during adolescence of the neural processing of social emotion. J. Cogn. Neurosci. *21*, 1736–1750. 10.1162/jocn.2009.21121.

33. Motta-Mena, N.V., and Scherf, K.S. (2017). Pubertal development shapes perception of complex facial expressions. Dev. Sci. *20*. 10.1111/desc.12451.

34. Glasser, M.F., Coalson, T.S., Robinson, E.C., Hacker, C.D., Harwell, J., Yacoub, E., Ugurbil, K., Andersson, J., Beckmann, C.F., Jenkinson, M., et al. (2016). A multi-modal parcellation of human cerebral cortex. Nature *536*, 171–178. 10.1038/nature18933.

35. Kaneshige, T., and Haryu, E. (2015). Categorization and understanding of facial expressions in 4-month-old infants. Jpn Psychol Res *57*, 135–142. 10.1111/jpr.12075.

36. Pollak, S.D., Messner, M., Kistler, D.J., and Cohn, J.F. (2009). Development of perceptual expertise in emotion recognition. Cognition *110*, 242–247. 10.1016/j.cognition.2008.10.010.

37. Ruba, A.L., and Pollak, S.D. (2020). The development of emotion reasoning in infancy and early childhood. Annu. Rev. Dev. Psychol. *2*, 503–531. 10.1146/annurev-devpsych-060320-102556.

38. Batty, M., and Taylor, M.J. (2006). The development of emotional face processing during childhood. Dev. Sci. *9*, 207–220. 10.1111/j.1467-7687.2006.00480.x.

39. Klasen, M., Chen, Y.-H., and Mathiak, K. (2012). Multisensory emotions: perception, combination and underlying neural processes. Rev. Neurosci. *23*, 381–392. 10.1515/revneuro-2012-0040.

40. Collignon, O., Girard, S., Gosselin, F., Roy, S., Saint-Amour, D., Lassonde, M., and Lepore, F. (2008). Audio-visual integration of emotion expression. Brain Res. *1242*, 126–135. 10.1016/j.brainres.2008.04.023.

41. Kreifelts, B., Ethofer, T., Grodd, W., Erb, M., and Wildgruber, D. (2007). Audiovisual integration of emotional signals in voice and face: an event-related fMRI study. Neuroimage *37*, 1445–1456. 10.1016/j.neuroimage.2007.06.020.

42. Weissman, D.H., Warner, L.M., and Woldorff, M.G. (2004). The neural mechanisms for minimizing cross-modal distraction. J. Neurosci. *24*, 10941–10949. 10.1523/JNEUROSCI.3669-04.2004.

43. Jacob, H., Brück, C., Domin, M., Lotze, M., and Wildgruber, D. (2014). I can't keep your face and voice out of my head: neural correlates of an attentional bias toward nonverbal emotional cues. Cereb. Cortex *24*, 1460–1473. 10.1093/cercor/bhs417.

44. Gunning-Dixon, F.M., Gur, R.C., Perkins, A.C., Schroeder, L., Turner, T., Turetsky, B.I., Chan, R.M., Loughead, J.W., Alsop, D.C., Maldjian, J., et al. (2003). Age-related differences in brain activation during emotional face processing. Neurobiol. Aging *24*, 285–295. 10.1016/s0197-4580(02)00099-4.

45. Monk, C.S., McClure, E.B., Nelson, E.E., Zarahn, E., Bilder, R.M., Leibenluft, E., Charney, D.S., Ernst, M., and Pine, D.S. (2003). Adolescent immaturity in attention-related brain engagement to emotional facial expressions. Neuroimage *20*, 420–428. 10.1016/s1053-8119(03)00355-0.

46. Lobaugh, N.J., Gibson, E., and Taylor, M.J. (2006). Children recruit distinct neural systems for implicit emotional face processing. Neuroreport *17*, 215–219. 10.1097/01.wnr.0000198946.00445.2f.