

Reviewed Preprint

v1 • August 7, 2025

Not revised

Reviewed Preprint

v2 • December 22, 2025

Revised by authors

For correspondence:

bijanki@bcm.edu

Competing interests:

No competing interests declared

Funding:

See [page 12](#)

Reviewing editor:

Xiaoqing Hu,
University of Hong Kong, China

© 2025, Fan et al. This article is distributed under the terms of the

Creative Commons Attribution

License, which permits unrestricted use and redistribution provided that the original author and source are credited.

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

Xiaoxu Fan¹, Abhishek Tripathi², Kelly R Bijanki¹ ✉

¹Baylor College of Medicine, Houston, United States • ²Rice University, Houston, United States

eLife Assessment

This study examines an **important** question regarding the developmental trajectory of neural mechanisms supporting facial expression processing. Leveraging a rare intracranial EEG (iEEG) dataset including both children and adults, the authors reported that facial expression recognition mainly engaged the posterior superior temporal cortex (pSTC) among children, while both pSTC and the prefrontal cortex were engaged among adults. In terms of strength of evidence, the **solid** methods, data and analyses broadly support the claims with minor weaknesses.

<https://doi.org/10.7554/eLife.107636.2.sa3>

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 area 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 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, young children rely more on low-level sensory area than on the prefrontal cortex for facial emotion processing, suggesting that the prefrontal cortex matures with development to enable a full understanding of facial emotions, especially complex emotions that require 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 (Barrera and Maurer, 1981 [DOI](#); Walker-Andrews, 1997 [DOI](#)), as infants just months old can distinguish happy and sad faces from surprised faces (Caron et al., 1982 [DOI](#); Nelson and Dolgin, 1985 [DOI](#); Nelson et al., 1979 [DOI](#)). However, children's emotion recognition is substantially less accurate than adults, and this ability prominently improves across childhood and adolescence (Johnston et al., 2011 [DOI](#); Kolb et al., 1992 [DOI](#); Lawrence et al., 2015 [DOI](#); Rodger et al., 2015 [DOI](#); Romani-Sponchiado et al., 2022 [DOI](#); Thomas et al., 2007 [DOI](#)). 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) (Duchaine and Yovel, 2015 [↗](#)). 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 (Gomez et al., 2017 [↗](#)). 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 (Arcaro et al., 2017 [↗](#)). It is likely that the gradual maturation of pSTS and FFA, two early sensory areas involved in the processing of facial expressions (Bernstein and Yovel, 2015 [↗](#); Engell and Haxby, 2007 [↗](#)), 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 (Pereira et al., 2019 [↗](#); Yurgelun-Todd and Killgore, 2006 [↗](#)). 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 (Kolb et al., 1992 [↗](#); Tessitore et al., 2005 [↗](#); Williams et al., 2006 [↗](#)). 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 (Hariri et al., 2003 [↗](#), 2000 [↗](#); Yurgelun-Todd, 2007 [↗](#)). However, 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 (Bernstein and Yovel, 2015 [↗](#); Duchaine and Yovel, 2015 [↗](#); Fan et al., 2020 [↗](#); Flack et al., 2015 [↗](#)). Furthermore, the encoding of complex emotions in the pSTC increases with age. These iEEG results 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.

Results

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

In this study, we analyzed iEEG data collected from a large group of epilepsy patients while they watched a short audiovisual film at the University Medical Center Utrecht (Berezutskaya et al., 2022 [↗](#)). 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 11 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 [↗](#)). We focused on HFB activity because it is widely considered to reflect the responses of local

neuronal populations near the recording electrode (Parvizi and Kastner, 2018 [↗](#)). The model performance was quantified as the correlation between the predicted and actual HFB activities, which is also called prediction accuracy.

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 [↗](#), gray bar, $P=0.012$, $t_{12}=2.954$, one-sample t 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 [↗](#), blue bar, $P=0.73$, $t_7=0.355$, one-sample t test). The prediction accuracy in children's DLPFC was significantly lower than in the post-childhood group ($P=0.036$, unpaired t test). Similar results were observed in the music condition. The prediction accuracy of the encoding model was significantly greater than zero in the post-childhood group (Figure 2C [↗](#), $P=0.012$, $t_{12}=2.975$, one-sample t test), but not in the childhood group (Figure 2C [↗](#), $P=0.19$, $t_7=1.469$, one-sample t test). Together, 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 2D [↗](#), $P=0.018$, unpaired t 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 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 (Bernstein and Yovel, 2015 [↗](#); Duchaine and Yovel, 2015 [↗](#); Fan et al., 2020 [↗](#); Flack et al., 2015 [↗](#)). Although previous studies suggested that the development of facial recognition depends on the maturation of face-selective brain regions (Arcaro et al., 2017 [↗](#); Gomez et al., 2017 [↗](#)), 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 children with electrode coverage in pSTC (Figure 3A [↗](#)). Our results showed that the encoding model significantly predicts the HFB neural signals in children's pSTC under the speech condition (Figure 3B [↗](#), $P<0.002$, two-tailed permutation test, see Method for details.) Moreover, the prediction accuracy is significantly reduced when human voice is absent from the video (Figure 3B [↗](#), $P=0.0306$, two-tailed permutation test). In an example subject (S19: 8 years old) who has electrodes in both DLPFC and pSTC, the model fitting results clearly showed that facial emotion is encoded in pSTC ($S19_{\text{speech}}$: $P=0.0014$, $r=0.1951$), but not DLPFC ($S19_{\text{speech}}$: $P=0.295$, $r=-0.0628$). Similarly, group-level results showed that the model performance is significantly greater than zero in the pSTC of post-childhood individuals (Figure 3D [↗](#) and 3E [↗](#), speech condition: $P=0.0001$, $t_{24}=4.601$, one-sample t test) and this neural representation of facial expression information is significantly reduced when human voice is absent (Figure 3E [↗](#), paired- t test, $t_{24}=2.897$, $P=0.0079$). These results provide evidence that children's sensory areas encode facial emotion features in a manner similar to that of post-childhood individuals.

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.

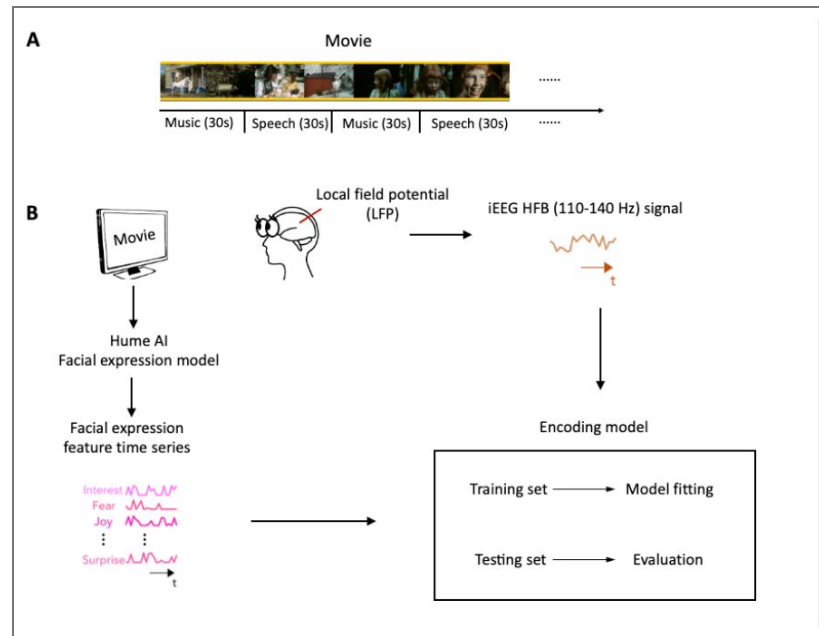
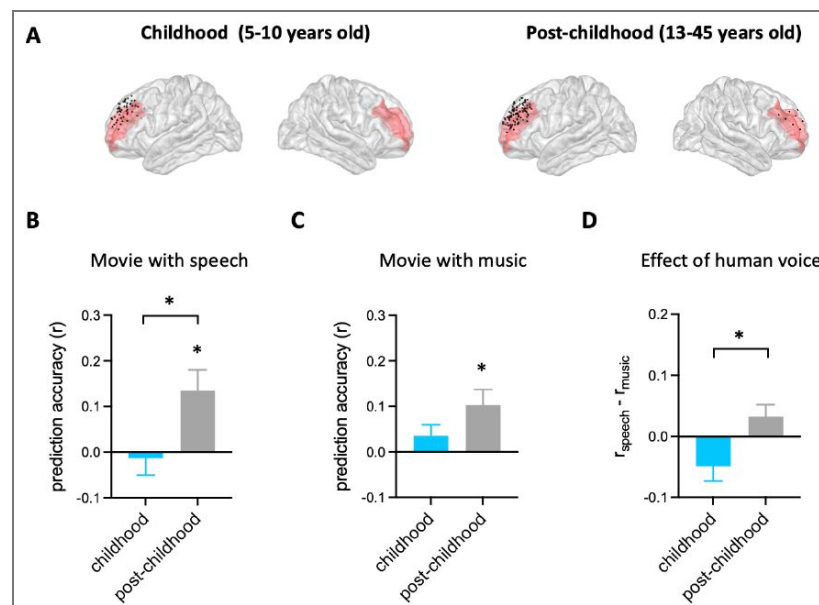


Figure 2. Prediction performance of encoding models in DLPFC.

(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 (Desikan et al., 2006). Electrodes outside DLPFC are not shown. (B) Averaged 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) Averaged prediction accuracy across participants for music condition. (D) Prediction accuracy difference between speech condition and music condition for each group. Error bars are standard error of the mean. $*P < 0.05$.



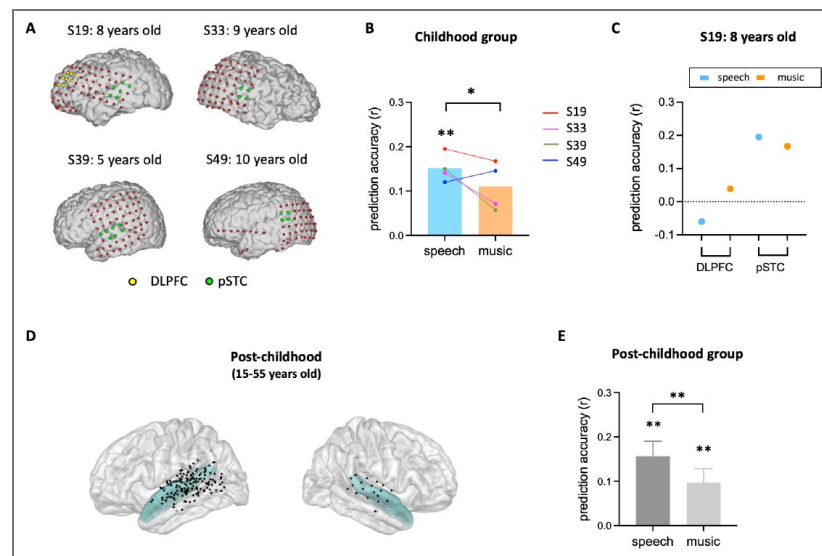


Figure 3. Prediction performance of encoding models in pSTC.

(A) The electrode distribution in the native brain space for the four children. Electrodes in pSTC are green, and electrodes in DLPFC are yellow. (B) Prediction accuracy of encoding models in children. Bars indicate mean prediction accuracy across participants and lines indicate individual data. Prediction accuracy of encoding models for S19 in both DLPFC and pSTC. (D) 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 (Desikan et al., 2006). (E) Averaged prediction accuracy across postchildhood participants. Error bars are standard error of the mean. ** $P < 0.01$. * $P < 0.05$.

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 (Alba-Ferrara et al., 2011; Burnett et al., 2011; Russell and Paris, 1994)) 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 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 (Burnett et al., 2009; Garcia and Scherf, 2015; Motta-Mena and Scherf, 2017).

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 (Caron et al., 1982; Nelson and Dolgin, 1985; Nelson et al., 1979). 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 (Kaneshige and Haryu, 2015) 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 (Pereira et al., 2019; Pollak et al., 2009; Ruba and Pollak, 2020). 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 (Batty and Taylor, 2006; Williams et al., 2006).

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

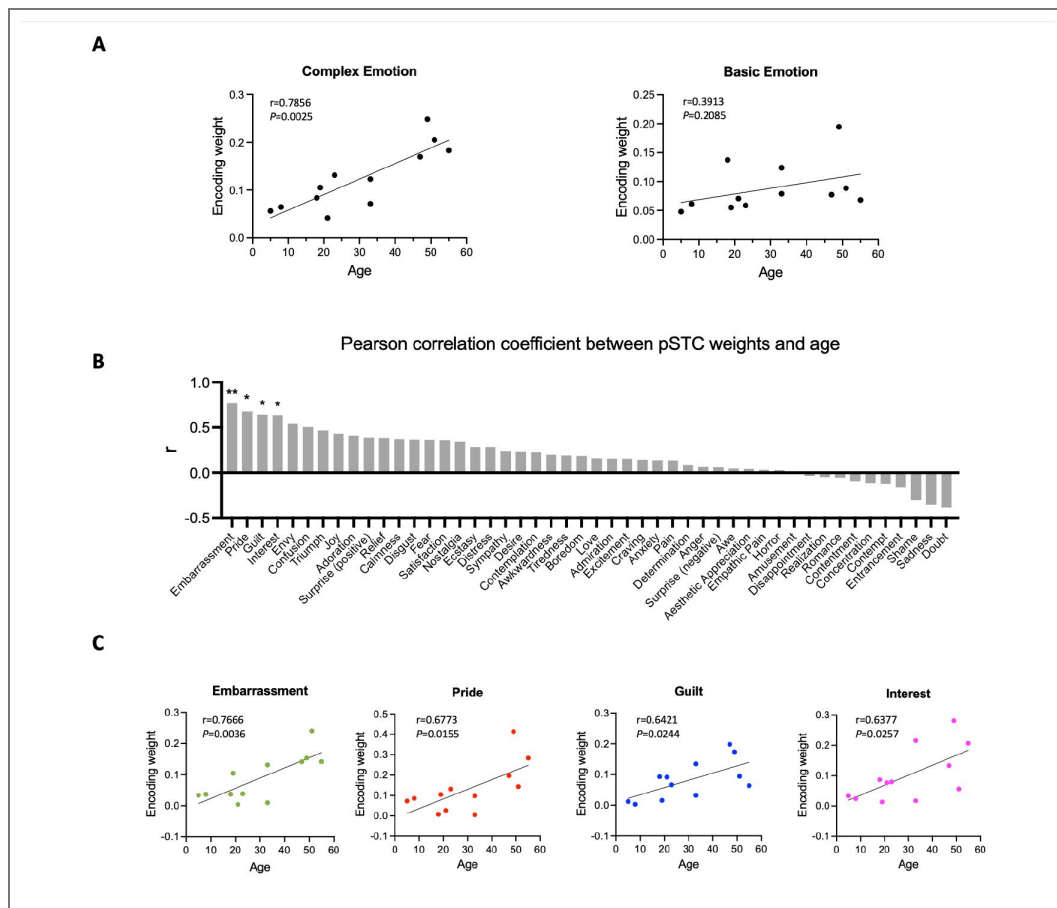


Figure 4. Correlation between encoding weights and age.

(A) Leb: 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 ($N=12$).

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 (Collignon et al., 2008 [↗](#); Klasen et al., 2012 [↗](#); Kreifelts et al., 2007 [↗](#)), whereas in children, the presence of voice might interfere with or redirect attentional resources away from facial expression processing (Jacob et al., 2014 [↗](#); Klasen et al., 2012 [↗](#); Weissman et al., 2004 [↗](#)).

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 (Gunning-Dixon et al., 2003 [↗](#); Tessitore et al., 2005 [↗](#); Thomas et al., 2007 [↗](#)), and for some emotion categories, this development may extend across the lifespan. For example, prefrontal cortex activation during viewing fearful faces increases with age (Monk et al., 2003 [↗](#); Williams et al., 2006 [↗](#); Yurgelun-Todd and Killgore, 2006 [↗](#)). 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 (Kolb et al., 1992 [↗](#)).

Our results clearly showed that facial expression is encoded in children's pSTC. Childhood and the post-childhood group showed comparable levels of prediction accuracy. In the 8-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 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 (Lobaugh et al., 2006 [↗](#)). Meanwhile, some other researchers found that responses to facial expression in the amygdala and posterior fusiform gyri decreased as people got older (Tessitore et al., 2005 [↗](#)), but the use of frontal regions increased with age (Gunning-Dixon et al., 2003 [↗](#)). 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, implying 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 (Burnett et al., 2009 [↗](#); Garcia and Scherf, 2015 [↗](#); Motta-Mena and Scherf, 2017 [↗](#)). In fact, our results implies 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.

Our study examines how facial emotions presented in naturalistic video stimuli are represented in the human brain across developmental stages. The facial emotional features used in our encoding models describe the emotions portrayed by characters in the stimulus, rather than participants'

subjective perception or experience of those emotions. While our results suggest the potential neural basis for participants' perception of these emotions, we cannot conclude that such information was consciously perceived without direct behavioral measures. In fact, automatic and efficient emotion perception is not driven by the structural features of a face alone but is also shaped by other factors, such as contextual information, prior experiences, and current mood states (Barrett et al., 2011 [↗](#)). How the conscious percept of facial emotion is constructed remains far from fully understood. Recent evidence suggests that dynamic interactions between visual cortices and frontal regions integrate stimulus-defined features with contextual and internal factors to give rise to subjective emotional perception (Wang et al., 2014 [↗](#)).

Limitations and future directions

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. Third, 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, or with longitudinal designs to directly track developmental trajectories, will be essential for refining our understanding of how facial emotion processing develops across the lifespan. In addition, future studies are needed to systematically examine how multimodal cues and prior experiences contribute to the understanding of emotion from faces.

Methods

In this study, iEEG data from an open multimodal iEEG-fMRI dataset were analyzed (Berezutskaya et al., 2022 [↗](#)).

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 [↗](#)). Eleven children (5-10 years old, 7 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 four 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. All patients were asked to participate the experiment outside the window of seizures. More data acquisition details can be found in Berezutskaya et al.'s article (Berezutskaya et al., 2022 [↗](#)).

Table 1. Demographic information of childhood group.

ID	Sex	Age	Number of recording contacts in left pSTC	Number of recording contacts in right pSTC	Number of recording contacts in left DLPFC	Number of recording contacts 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
s33	F	9	0	4	0	0
s37	M	5	0	0	4	0
s39	F	5	7	0	0	0
s41	M	7	0	0	4	0
s49	F	10	4	0	0	0
s50	F	9	0	0	6	0
s63	M	5	0	0	9	0

Table 2. Demographic information of post-childhood group.

ID	Sex	Age	Number of recording contacts in left pSTC	Number of recording contacts in right pSTC	Number of recording contacts in left DLPFC	Number of recording contacts 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	6	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	6	0	0	0
s59	F	30	7	0	0	0
s60	M	42	8	0	0	0
s61	F	16	8	0	0	0

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 (Glasser et al., 2016) 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 2 Hz. 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.

One-sample t test was used to test whether the encoding model performance was significantly > 0. For children's pSTC, 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.

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.

Data availability

iEEG data is available in openneuro. Source data for Figure 2-4 will be available in OSF.

Acknowledgements

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).

Additional information

Author contributions

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

Funding

Funder	Grant reference number	Author
HHS NIH National Institute of Mental Health (NIMH)		Kelly Bijanki

Author ORCID iDs

Xiaoxu Fan: <https://orcid.org/0000-0002-8115-8621>

Kelly R Bijanki: <https://orcid.org/0000-0003-1624-8767>

References

- Alba-Ferrara L**, Hausmann M, Mitchell RL, Weis S. (2011) The neural correlates of emotional prosody comprehension: disentangling simple from complex emotion. *PLoS ONE* **6**:e28701 <https://doi.org/10.1371/journal.pone.0028701>
- Arcaro MJ**, Schade PF, Vincent JL, Ponce CR, Livingstone MS (2017) Seeing faces is necessary for facedomain formation. *Nat Neurosci* **20**:1404-1412 <https://doi.org/10.1038/nn.4635>
- Barrera ME**, Maurer D. (1981) The perception of facial expressions by the three-month-old. *Child Dev* **52**:203-206 <https://doi.org/10.2307/1129231>
- Barrett LF**, Mesquita B, Gendron M. (2011) Context in Emotion Perception. *Current Directions in Psychological Science* **20**:286-290 <https://doi.org/10.1177/0963721411422522>
- Batty M**, Taylor MJ (2006) The development of emotional face processing during childhood. *Dev Sci* **9**:207-220 <https://doi.org/10.1111/j.1467-7687.2006.00480.x>

- Berezutskaya J**, Vansteensel MJ, Aarnoutse EJ, Freudenburg ZV, Piantoni G, Branco MP, Ramsey NF (2022) Open multimodal iEEG-fMRI dataset from naturalistic stimulation with a short audiovisual film. *Sci Data* **9**:91 <https://doi.org/10.1038/s41597-022-01173-0>
- Bernstein M**, Yovel G. (2015) Two neural pathways of face processing: A critical evaluation of current models. *Neurosci Biobehav Rev* **55**:536-546 <https://doi.org/10.1016/j.neubiorev.2015.06.010>
- Burnett S**, Bird G, Moll J, Frith C, Blakemore S-J. (2009) Development during adolescence of the neural processing of social emotion. *J Cogn Neurosci* **21**:1736-1750 <https://doi.org/10.1162/jocn.2009.21121>
- Burnett S**, Thompson S, Bird G, Blakemore S-J. (2011) Pubertal development of the understanding of social emotions: Implications for education. *Learn Individ Differ* **21**:681-689 <https://doi.org/10.1016/j.lindif.2010.05.007>
- Caron RF**, Caron AJ, Myers RS (1982) Abstraction of invariant face expressions in infancy. *Child Dev* **53**:1008-1015
- Collignon O**, Girard S, Gosselin F, Roy S, Saint-Amour D, Lassonde M, Lepore F. (2008) Audio-visual integration of emotion expression. *Brain Res* **1242**:126-135 <https://doi.org/10.1016/j.brainres.2008.04.023>
- Desikan RS**, Ségonne F, Fischl B, Quinn BT, Dickerson BC, Blacker D, Buckner RL, Dale AM, Maguire RP, Hyman BT, *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 <https://doi.org/10.1016/j.neuroimage.2006.01.021>
- Duchaine B**, Yovel G. (2015) A revised neural framework for face processing. *Annu Rev Vis Sci* **1**:393-416 <https://doi.org/10.1146/annurev-vision-082114-035518>
- Engell AD**, Haxby JV (2007) Facial expression and gaze-direction in human superior temporal sulcus. *Neuropsychologia* **45**:3234-3241 <https://doi.org/10.1016/j.neuropsychologia.2007.06.022>
- Fan X**, Wang F, Shao H, Zhang P, He S. (2020) The bottom-up and top-down processing of faces in the human occipitotemporal cortex. *eLife* **9** <https://doi.org/10.7554/eLife.48764>
- Flack TR**, Andrews TJ, Hymers M, Al-Mosaiwi M, Marsden SP, Strachan JWA, Trakulpipat C, Wang L, Wu T, Young AW (2015) Responses in the right posterior superior temporal sulcus show a feature-based response to facial expression. *Cortex* **69**:14-23 <https://doi.org/10.1016/j.cortex.2015.03.002>
- Garcia NV**, Scherf KS (2015) Emerging sensitivity to socially complex expressions: A unique role for adolescence?. *Child Dev Perspect* **9**:84-90 <https://doi.org/10.1111/cdep.12114>
- Glasser MF**, Coalson TS, Robinson EC, Hacker CD, Harwell J, Yacoub E, Ugurbil K, Andersson J, Beckmann CF, Jenkinson M, *et al.* (2016) A multi-modal parcellation of human cerebral cortex. *Nature* **536**:171-178 <https://doi.org/10.1038/nature18933>
- Gomez J**, Barnett MA, Natu V, Mezer A, Palomero-Gallagher N, Weiner KS, Amunts K, Zilles K, Grill-Spector K. (2017) Microstructural proliferation in human cortex is coupled with the development of face processing. *Science* **355**:68-71 <https://doi.org/10.1126/science.aag0311>
- Gunning-Dixon FM**, Gur RC, Perkins AC, Schroeder L, Turner T, Turetsky BI, Chan RM, Loughhead JW, Alsop DC, Maldjian J, *et al.* (2003) Age-related differences in brain activation during emotional face processing. *Neurobiol Aging* **24**:285-295 [https://doi.org/10.1016/s0197-4580\(02\)00099-4](https://doi.org/10.1016/s0197-4580(02)00099-4)
- Hariri AR**, Bookheimer SY, Mazziotta JC (2000) Modulating emotional responses: effects of a neocortical network on the limbic system. *Neuroreport* **11**:43-48 <https://doi.org/10.1097/00001756-200001170-00009>
- Hariri AR**, Mattay VS, Tessitore A, Fera F, Weinberger DR (2003) Neocortical modulation of the amygdala response to fearful stimuli. *Biol Psychiatry* **53**:494-501 [https://doi.org/10.1016/s0006-3223\(02\)01786-9](https://doi.org/10.1016/s0006-3223(02)01786-9)
- Jacob H**, Brück C, Domin M, Lotze M, 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 <https://doi.org/10.1093/cercor/bhs417>

- Johnston PJ, Kaufman J, Bajic J, Sercombe A, Michie PT, Karayanidis F. (2011) Facial emotion and identity processing development in 5- to 15-year-old children. *Front Psychol* **2**:26 <https://doi.org/10.3389/fpsyg.2011.00026>
- Kaneshige T, Haryu E. (2015) Categorization and understanding of facial expressions in 4-month-old infants. *Jpn Psychol Res* **57**:135-142 <https://doi.org/10.1111/jpr.12075>
- Klasen M, Chen Y-H, Mathiak K. (2012) Multisensory emotions: perception, combination and underlying neural processes. *Rev Neurosci* **23**:381-392 <https://doi.org/10.1515/revneuro-2012-0040>
- Kolb B, Wilson B, Taylor L. (1992) Developmental changes in the recognition and comprehension of facial expression: implications for frontal lobe function. *Brain Cogn* **20**:74-84 [https://doi.org/10.1016/0278-2626\(92\)90062-q](https://doi.org/10.1016/0278-2626(92)90062-q)
- Kreifelts B, Ethofer T, Grodd W, Erb M, Wildgruber D. (2007) Audiovisual integration of emotional signals in voice and face: an event-related fMRI study. *Neuroimage* **37**:1445-1456 <https://doi.org/10.1016/j.neuroimage.2007.06.020>
- Lawrence K, Campbell R, Skuse D. (2015) Age, gender, and puberty influence the development of facial emotion recognition. *Front Psychol* **6**:761 <https://doi.org/10.3389/fpsyg.2015.00761>
- Lobaugh NJ, Gibson E, Taylor MJ (2006) Children recruit distinct neural systems for implicit emotional face processing. *Neuroreport* **17**:215-219 <https://doi.org/10.1097/01.wnr.0000198946.00445.2f>
- Monk CS, McClure EB, Nelson EE, Zarah E, Bilder RM, Leibenluft E, Charney DS, Ernst M, Pine DS (2003) Adolescent immaturity in attention-related brain engagement to emotional facial expressions. *Neuroimage* **20**:420-428 [https://doi.org/10.1016/s1053-8119\(03\)00355-0](https://doi.org/10.1016/s1053-8119(03)00355-0)
- Motta-Mena NV, Scherf KS (2017) Pubertal development shapes perception of complex facial expressions. *Dev Sci* **20** <https://doi.org/10.1111/desc.12451>
- Nelson CA, Dolgin KG (1985) The generalized discrimination of facial expressions by seven-month-old infants. *Child Dev* **56**:58-61
- Nelson CA, Morse PA, Leavitt LA (1979) Recognition of facial expressions by seven-month-old infants. *Child Dev* **50**:1239-1242 <https://doi.org/10.2307/1129358>
- Parvizi J, Kastner S. (2018) Promises and limitations of human intracranial electroencephalography. *Nat Neurosci* **21**:474-483 <https://doi.org/10.1038/s41593-018-0108-2>
- Pereira MR, Barbosa F, de Haan M, Ferreira-Santos F. (2019) Understanding the development of face and emotion processing under a predictive processing framework. *Dev Psychol* **55**:1868-1881 <https://doi.org/10.1037/dev0000706>
- Pollak SD, Messner M, Kistler DJ, Cohn JF (2009) Development of perceptual expertise in emotion recognition. *Cognition* **110**:242-247 <https://doi.org/10.1016/j.cognition.2008.10.010>
- Rodger H, Vizioli L, Ouyang X, Caldara R. (2015) Mapping the development of facial expression recognition. *Dev Sci* **18**:926-939 <https://doi.org/10.1111/desc.12281>
- Romani-Sponchiado A, Maia CP, Torres CN, Tavares I, Arteché AX (2022) Emotional face expressions recognition in childhood: developmental markers, age and sex effect. *Cogn Process* **23**:467-477 <https://doi.org/10.1007/s10339-022-01086-1>
- Ruba AL, Pollak SD (2020) The development of emotion reasoning in infancy and early childhood. *Annu Rev Dev Psychol* **2**:503-531 <https://doi.org/10.1146/annurev-devpsych-060320-102556>
- Russell JA, Paris FA (1994) Do Children acquire Concepts for Complex Emotions Abruptly?. *Int J Behav Dev* **17**:349-365 <https://doi.org/10.1177/016502549401700207>
- Tessitore A, Hariri AR, Fera F, Smith WG, Das S, Weinberger DR, Mattay VS (2005) Functional changes in the activity of brain regions underlying emotion processing in the elderly. *Psychiatry Res* **139**:9-18 <https://doi.org/10.1016/j.psychres.2005.02.009>
- Thomas LA, De Bellis MD, Graham R, LaBar KS (2007) Development of emotional facial recognition in late childhood and adolescence. *Dev Sci* **10**:547-558 <https://doi.org/10.1111/j.1467-7687.2007.00614.x>

- Walker-Andrews AS (1997) Infants' perception of expressive behaviors: differentiation of multimodal information. *Psychol Bull* **121**:437-456 <https://doi.org/10.1037/0033-2909.121.3.437>
- Wang S, Tudusciuc O, Mamelak AN, Ross IB, Adolphs R, Rutishauser U. (2014) Neurons in the human amygdala selective for perceived emotion. *Proc Natl Acad Sci USA* **111**:E3110-9 <https://doi.org/10.1073/pnas.1323342111>
- Weissman DH, Warner LM, Woldorff MG (2004) The neural mechanisms for minimizing cross-modal distraction. *J Neurosci* **24**:10941-10949 <https://doi.org/10.1523/JNEUROSCI.3669-04.2004>
- Williams LM, Brown KJ, Palmer D, Liddell BJ, Kemp AH, Olivieri G, Peduto A, Gordon E. (2006) The mellow years?: neural basis of improving emotional stability over age. *J Neurosci* **26**:6422-6430 <https://doi.org/10.1523/JNEUROSCI.0022-06.2006>
- Yurgelun-Todd DA, Killgore WDS (2006) Fear-related activity in the prefrontal cortex increases with age during adolescence: a preliminary fMRI study. *Neurosci Lett* **406**:194-199 <https://doi.org/10.1016/j.neulet.2006.07.046>
- Yurgelun-Todd D. (2007) Emotional and cognitive changes during adolescence. *Curr Opin Neurobiol* **17**:251-257 <https://doi.org/10.1016/j.conb.2007.03.009>
- Julia Berezutskaya, Mariska J. Vansteensel, Erik J. Aarnoutse, Zachary V. Freudenburg, Giovanni Piantoni, Mariana P. Branco, Nick F. Ramsey (2022) Open multimodal iEEG-fMRI dataset from naturalistic stimulation with a short audiovisual film. OpenNeuro. ID ds003688 <https://openneuro.org/datasets/ds003688/versions/1.0.7>

Peer reviews

Reviewer #1 (Public review):

Summary:

This study investigates how the brain processes facial expressions across development by analyzing intracranial EEG (iEEG) data from children (ages 5-10) and post-childhood individuals (ages 13-55). The researchers used a short film containing emotional facial expressions and applied AI-based models to decode brain responses to facial emotions. They found that in children, facial emotion information is represented primarily in the posterior superior temporal cortex (pSTC)-a sensory processing area-but not in the dorsolateral prefrontal cortex (DLPFC), which is involved in higher-level social cognition. In contrast, post-childhood individuals showed emotion encoding in both regions. Importantly, the complexity of emotions encoded in the pSTC increased with age, particularly for socially nuanced emotions like embarrassment, guilt, and pride. The authors claim that these findings suggest that emotion recognition matures through increasing involvement of the prefrontal cortex, supporting a developmental trajectory where top-down modulation enhances understanding of complex emotions as children grow older.

Strengths:

- (1) The inclusion of pediatric iEEG makes this study uniquely positioned to offer high-resolution temporal and spatial insights into neural development compared to non-invasive approaches, e.g., fMRI, scalp EEG, etc.
- (2) Using a naturalistic film paradigm enhances ecological validity compared to static image tasks often used in emotion studies.
- (3) The idea of using state-of-the-art AI models to extract facial emotion features allows for high-dimensional and dynamic emotion labeling in real time.

Weaknesses:

- (1) The study has notable limitations that constrain the generalizability and depth of its conclusions. The sample size was very small, with only nine children included and just two having sufficient electrode coverage in the posterior superior temporal cortex (pSTC), which weakens the reliability and statistical power of the findings, especially for analyses involving age. Authors pointed out that a similar sample size has been used in previous iEEG studies, but the cited works focus on adults and do not look at the developmental perspectives. Similar work looking at developmental changes in iEEG signals usually includes many more subjects (e.g., $n = 101$ children from Cross ZR et al., *Nature Human Behavior*, 2025) to account for inter-subject variabilities.
- (2) Electrode coverage was also uneven across brain regions, with not all participants having electrodes in both the dorsolateral prefrontal cortex (DLPFC) and pSTC, making the conclusion regarding the different developmental changes between DLPFC and pSTC hard to interpret (related to point 3 below). It is understood that it is rare to have such iEEG data collected in this age group, and the electrode location is only determined by clinical needs. However, the scientific rigor should not be compromised by the limited data access. It's the authors' decision whether such an approach is valid and appropriate to address the scientific questions, here the developmental changes in the brain, given all the advantages and constraints of the data modality.
- (3) The developmental differences observed were based on cross-sectional comparisons rather than longitudinal data, reducing the ability to draw causal conclusions about developmental trajectories. Also, see comments in point 2.
- (4) Moreover, the analysis focused narrowly on DLPFC, neglecting other relevant prefrontal areas such as the orbitofrontal cortex (OFC) and anterior cingulate cortex (ACC), which play key roles in emotion and social processing. Agree that this might be beyond the scope of this paper, but a discussion section might be insightful.
- (5) Although the use of a naturalistic film stimulus enhances ecological validity, it comes at the cost of experimental control, with no behavioral confirmation of the emotions perceived by participants and uncertain model validity for complex emotional expressions in children. A non-facial music block that could have served as a control was available but not analyzed. The validation of AI model's emotional output needs to be tested. It is understood that we cannot collect these behavioral data retrospectively within the recorded subjects. Maybe potential post-hoc experiments and analyses could be done, e.g., collect behavioral, emotional perception data from age-matched healthy subjects.
- (6) Generalizability is further limited by the fact that all participants were neurosurgical patients, potentially with neurological conditions such as epilepsy that may influence brain responses. At least some behavioral measures between the patient population and the healthy groups should be done to ensure the perception of emotions is similar.
- (7) Additionally, the high temporal resolution of intracranial EEG was not fully utilized, as data were downsampled and averaged in 500-ms windows. It seems like the authors are trying to compromise the iEEG data analyses to match up with the AI's output resolution, which is 2Hz. It is not clear then why not directly use fMRI, which is non-invasive and seems to meet the needs here already. The advantages of using iEEG in this study are missing here.
- (8) Finally, the absence of behavioral measures or eye-tracking data makes it difficult to directly link neural activity to emotional understanding or determine which facial features participants attended to. Related to point 5 as well.

Comments on revisions:

A behavioral measurement will help address a lot of these questions. If the data continues collecting, additional subjects with iEEG recording and also behavioral measurements would be valuable.

<https://doi.org/10.7554/eLife.107636.2.sa2>

Reviewer #2 (Public review):

Summary:

In this paper, Fan et al. aim to characterize how neural representations of facial emotions evolve from childhood to adulthood. Using intracranial EEG recordings from participants aged 5 to 55, the authors assess the encoding of emotional content in high-level cortical regions. They report that while both the posterior superior temporal cortex (pSTC) and dorsolateral prefrontal cortex (DLPFC) are involved in representing facial emotions in older individuals, only the pSTC shows significant encoding in children. Moreover, the encoding of complex emotions in the pSTC appears to strengthen with age. These findings lead the authors to suggest that young children rely more on low-level sensory areas and propose a developmental shift from reliance on lower-level sensory areas in early childhood to increased top-down modulation by the prefrontal cortex as individuals mature.

Strengths:

- (1) Rare and valuable dataset: The use of intracranial EEG recordings in a developmental sample is highly unusual and provides a unique opportunity to investigate neural dynamics with both high spatial and temporal resolution.
- (2) Developmentally relevant design: The broad age range and cross-sectional design are well-suited to explore age-related changes in neural representations.
- (3) Ecological validity: The use of naturalistic stimuli (movie clips) increases the ecological relevance of the findings.
- (4) Feature-based analysis: The authors employ AI-based tools to extract emotion-related features from naturalistic stimuli, which enables a data-driven approach to decoding neural representations of emotional content. This method allows for a more fine-grained analysis of emotion processing beyond traditional categorical labels.

Weaknesses:

- (1) While the authors leverage Hume AI, a tool pre-trained on a large dataset, its specific performance on the stimuli used in this study remains unverified. To strengthen the foundation of the analysis, it would be important to confirm that Hume AI's emotional classifications align with human perception for these particular videos. A straightforward way to address this would be to recruit human raters to evaluate the emotional content of the stimuli and compare their ratings to the model's outputs.
- (2) Although the study includes data from four children with pSTC coverage—an increase from the initial submission—the sample size remains modest compared to recent iEEG studies in the field.
- (3) The "post-childhood" group (ages 13-55) conflates several distinct neurodevelopmental periods, including adolescence, young adulthood, and middle adulthood. As a finer age stratification is likely not feasible with the current sample size, I would suggest authors temper their developmental conclusions.

(4) The analysis of DLPFC-pSTC directional connectivity would be significantly strengthened by modeling it as a continuous function of age across all participants, rather than relying on an unbalanced comparison between a single child and a (N=7) post-childhood group. This continuous approach would provide a more powerful and nuanced view of the developmental trajectory. I would also suggest including the result in the main text.

<https://doi.org/10.7554/eLife.107636.2.sa1>

Author response:

The following is the authors' response to the original reviews.

eLife Assessment

This study examines a valuable question regarding the developmental trajectory of neural mechanisms supporting facial expression processing. Leveraging a rare intracranial EEG (iEEG) dataset including both children and adults, the authors reported that facial expression recognition mainly engaged the posterior superior temporal cortex (pSTC) among children, while both pSTC and the prefrontal cortex were engaged among adults. However, the sample size is relatively small, with analyses appearing incomplete to fully support the primary claims.

Public Reviews:

Reviewer #1 (Public review):

Summary:

This study investigates how the brain processes facial expressions across development by analyzing intracranial EEG (iEEG) data from children (ages 5-10) and post-childhood individuals (ages 13-55). The researchers used a short film containing emotional facial expressions and applied AI-based models to decode brain responses to facial emotions. They found that in children, facial emotion information is represented primarily in the posterior superior temporal cortex (pSTC) - a sensory processing area - but not in the dorsolateral prefrontal cortex (DLPFC), which is involved in higher-level social cognition. In contrast, post-childhood individuals showed emotion encoding in both regions. Importantly, the complexity of emotions encoded in the pSTC increased with age, particularly for socially nuanced emotions like embarrassment, guilt, and pride. The authors claim that these findings suggest that emotion recognition matures through increasing involvement of the prefrontal cortex, supporting a developmental trajectory where top-down modulation enhances understanding of complex emotions as children grow older.

Strengths:

(1) The inclusion of pediatric iEEG makes this study uniquely positioned to offer high-resolution temporal and spatial insights into neural development compared to non-invasive approaches, e.g., fMRI, scalp EEG, etc.

(2) Using a naturalistic film paradigm enhances ecological validity compared to static image tasks often used in emotion studies.

(3) The idea of using state-of-the-art AI models to extract facial emotion features allows for high-dimensional and dynamic emotion labeling in real time

Weaknesses:

(1) The study has notable limitations that constrain the generalizability and depth of its conclusions. The sample size was very small, with only nine children included and just two having sufficient electrode coverage in the posterior superior temporal cortex (pSTC), which weakens the reliability and statistical power of the findings, especially for analyses involving age

We appreciated the reviewer's point regarding the constrained sample size.

As an invasive method, iEEG recordings can only be obtained from patients undergoing electrode implantation for clinical purposes. Thus, iEEG data from young children are extremely rare, and rapidly increasing the sample size within a few years is not feasible. However, we are confident in the reliability of our main conclusions. Specifically, 8 children (53 recording contacts in total) and 13 control participants (99 recording contacts in total) with electrode coverage in the DLPFC are included in our DLPFC analysis. This sample size is comparable to other iEEG studies with similar experiment designs [1-3].

For pSTC, we returned to the data set and found another two children who had pSTC coverage. After involving these children's data, the group-level analysis using permutation test showed that children's pSTC significantly encode facial emotion in naturalistic contexts (Figure 3B). Notably, the two new children's (S33 and S49) responses were highly consistent with our previous observations. Moreover, the averaged prediction accuracy in children's pSTC ($r_{\text{speech}}=0.1565$) was highly comparable to that in post-childhood group ($r_{\text{speech}}=0.1515$).

(1) Zheng, J. et al. Multiplexing of Theta and Alpha Rhythms in the Amygdala-Hippocampal Circuit Supports Pafern Separation of Emotional Information. *Neuron* 102, 887-898.e5 (2019).

(2) Diamond, J. M. et al. Focal seizures induce spatiotemporally organized spiking activity in the human cortex. *Nat. Commun.* 15, 7075 (2024).

(3) Schrouff, J. et al. Fast temporal dynamics and causal relevance of face processing in the human temporal cortex. *Nat. Commun.* 11, 656 (2020).

(2) Electrode coverage was also uneven across brain regions, with not all participants having electrodes in both the dorsolateral prefrontal cortex (DLPFC) and pSTC, and most coverage limited to the left hemisphere-hindering within-subject comparisons and limiting insights into lateralization.

The electrode coverage in each patient is determined entirely by the clinical needs. Only a few patients have electrodes in both DLPFC and pSTC because these two regions are far apart, so it's rare for a single patient's suspected seizure network to span such a large territory. However, it does not affect our results, as most iEEG studies combine data from multiple patients to achieve sufficient electrode coverage in each target brain area. As our data are mainly from left hemisphere (due to the clinical needs), this study was not designed to examine whether there is a difference between hemispheres in emotion encoding. Nevertheless, lateralization remains an interesting question that should be addressed in future research, and we have noted this limitation in the Discussion (Page 8, in the last paragraph of the Discussion).

(3) The developmental differences observed were based on cross-sectional comparisons rather than longitudinal data, reducing the ability to draw causal conclusions about developmental trajectories.

In the context of pediatric intracranial EEG, longitudinal data collection is not feasible due to the invasive nature of electrode implantation. We have added this point to the Discussion to acknowledge that while our results reveal robust age-related differences in the cortical encoding of facial emotions, longitudinal studies using non-invasive methods will be essential

to directly track developmental trajectories (Page 8, in the last paragraph of Discussion). In addition, we revised our manuscript to avoid emphasis causal conclusions about developmental trajectories in the current study (For example, we use “imply” instead of “suggest” in the fifth paragraph of Discussion).

(4) Moreover, the analysis focused narrowly on DLPFC, neglecting other relevant prefrontal areas such as the orbitofrontal cortex (OFC) and anterior cingulate cortex (ACC), which play key roles in emotion and social processing.

We agree that both OFC and ACC are critically involved in emotion and social processing. However, we have no recordings from these areas because ECoG rarely covers the ACC or OFC due to technical constraints. We have noted this limitation in the Discussion (Page 8, in the last paragraph of Discussion). Future follow-up studies using sEEG or non-invasive imaging methods could be used to examine developmental patterns in these regions.

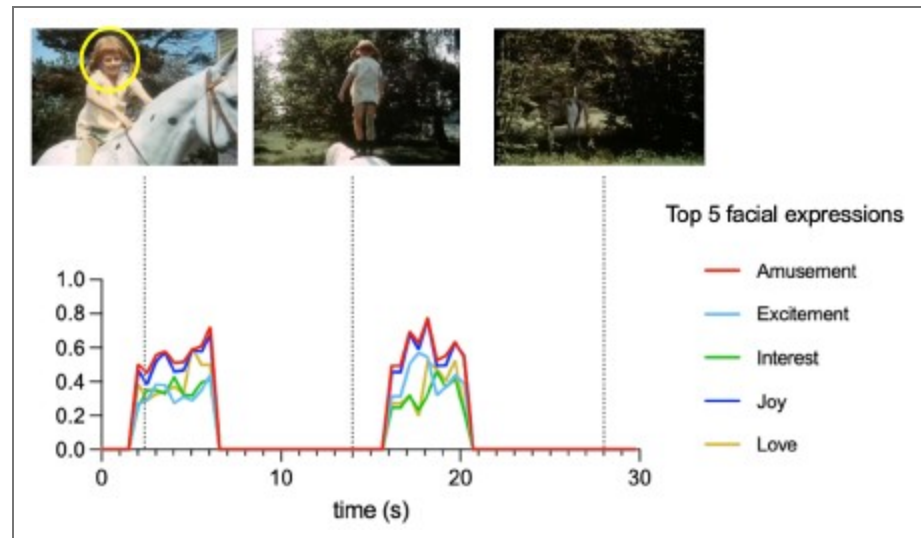
(5) Although the use of a naturalistic film stimulus enhances ecological validity, it comes at the cost of experimental control, with no behavioral confirmation of the emotions perceived by participants and uncertain model validity for complex emotional expressions in children. A nonfacial music block that could have served as a control was available but not analyzed.

The facial emotion features used in our encoding models were extracted by Hume AI models, which were trained on human intensity ratings of large-scale, experimentally controlled emotional expression data[1-2]. Thus, the outputs of Hume AI model reflect what typical facial expressions convey, that is, the presented facial emotion. Our goal of the present study was to examine how facial emotions presented in the videos are encoded in the human brain at different developmental stages. We agree that children’s interpretation of complex emotions may differ from that of adults, resulting in different perceived emotion (i.e., the emotion that the observer subjectively interprets). Behavioral ratings are necessary to study the encoding of subjectively perceived emotion, which is a very interesting direction but beyond the scope of the present work. We have added a paragraph in the Discussion (see Page 8) to explicitly note that our study focused on the encoding of presented emotion.

We appreciated the reviewer’s point regarding the value of non-facial music blocks. However, although there are segments in music condition that have no faces presented, these cannot be used as a control condition to test whether the encoding model’s prediction accuracy in pSTC or DLPFC drops to chance when no facial emotion is present. This is because, in the absence of faces, no extracted emotion features are available to be used for the construction of encoding model (see Author response image 1 below). Thus, we chose to use a different control analysis for the present work. For children’s pSTC, we shuffled facial emotion feature in time to generate a null distribution, which was then used to test the statistical significance of the encoding models (see Methods/Encoding model fitting for details).

(1) Brooks, J. A. et al. Deep learning reveals what facial expressions mean to people in different cultures. *iScience* 27, 109175 (2024).

(2) Brooks, J. A. et al. Deep learning reveals what vocal bursts express in different cultures. *Nat. Hum. Behav.* 7, 240–250 (2023).



Author response image 1. Time courses of Hume AI extracted facial expression features for the first block of music condition. Only top 5 facial expressions were shown here to due to space limitation.

(6) Generalizability is further limited by the fact that all participants were neurosurgical patients, potentially with neurological conditions such as epilepsy that may influence brain responses.

We appreciated the reviewer's point. However, iEEG data can only be obtained from clinical populations (usually epilepsy patients) who have electrodes implantation. Given current knowledge about focal epilepsy and its potential effects on brain activity, researchers believe that epilepsy-affected brains can serve as a reasonable proxy for normal human brains when confounding influences are minimized through rigorous procedures[1]. In our study, we took several steps to ensure data quality: (1) all data segments containing epileptiform discharges were identified and removed at the very beginning of preprocessing, (2) patients were asked to participate the experiment several hours outside the window of seizures. Please see Method for data quality check description (Page 9/ Experimental procedures and iEEG data processing).

(1) Parvizi J, Kastner S. 2018. Promises and limitations of human intracranial electroencephalography. *Nat Neurosci* 21:474–483. doi:10.1038/s41593-018-0108-2

(7) Additionally, the high temporal resolution of intracranial EEG was not fully utilized, as data were down-sampled and averaged in 500-ms windows.

We agree that one of the major advantages of iEEG is its millisecond-level temporal resolution. In our case, the main reason for down-sampling was that the time series of facial emotion features extracted from the videos had a temporal resolution of 2 Hz, which were used for the modelling neural responses. In naturalistic contexts, facial emotion features do not change on a millisecond timescale, so a 500 ms window is sufficient to capture the relevant dynamics. Another advantage of iEEG is its tolerance to motion, which is excessive in young children (e.g., 5-year-olds). This makes our dataset uniquely valuable, suggesting robust representation in the pSTC but not in the DLPFC in young children. Moreover, since our method framework (Figure 1) does not rely on high temporal resolution method, so it can be transferred to non-invasive modalities such as fMRI, enabling future studies to test these developmental patterns in larger populations.

(8) Finally, the absence of behavioral measures or eye-tracking data makes it difficult to directly link neural activity to emotional understanding or determine which facial features participants attended to.

We appreciated this point. Part of our rationale is presented in our response to (5) for the absence of behavioral measures. Following the same rationale, identifying which facial features participants attended to is not necessary for testing our main hypotheses because our analyses examined responses to the overall emotional content of the faces. However, we agree and recommend future studies use eye-tracking and corresponding behavioral measures in studies of subjective emotional understanding.

Reviewer #2 (Public review):

Summary:

In this paper, Fan et al. aim to characterize how neural representations of facial emotions evolve from childhood to adulthood. Using intracranial EEG recordings from participants aged 5 to 55, the authors assess the encoding of emotional content in high-level cortical regions. They report that while both the posterior superior temporal cortex (pSTC) and dorsolateral prefrontal cortex (DLPFC) are involved in representing facial emotions in older individuals, only the pSTC shows significant encoding in children. Moreover, the encoding of complex emotions in the pSTC appears to strengthen with age. These findings lead the authors to suggest that young children rely more on low-level sensory areas and propose a developmental shift from reliance on lower-level sensory areas in early childhood to increased top-down modulation by the prefrontal cortex as individuals mature.

Strengths:

(1) Rare and valuable dataset: The use of intracranial EEG recordings in a developmental sample is highly unusual and provides a unique opportunity to investigate neural dynamics with both high spatial and temporal resolution.

(2) Developmentally relevant design: The broad age range and cross-sectional design are well-suited to explore age-related changes in neural representations.

(3) Ecological validity: The use of naturalistic stimuli (movie clips) increases the ecological relevance of the findings.

(4) Feature-based analysis: The authors employ AI-based tools to extract emotion-related features from naturalistic stimuli, which enables a data-driven approach to decoding neural representations of emotional content. This method allows for a more fine-grained analysis of emotion processing beyond traditional categorical labels.

Weaknesses:

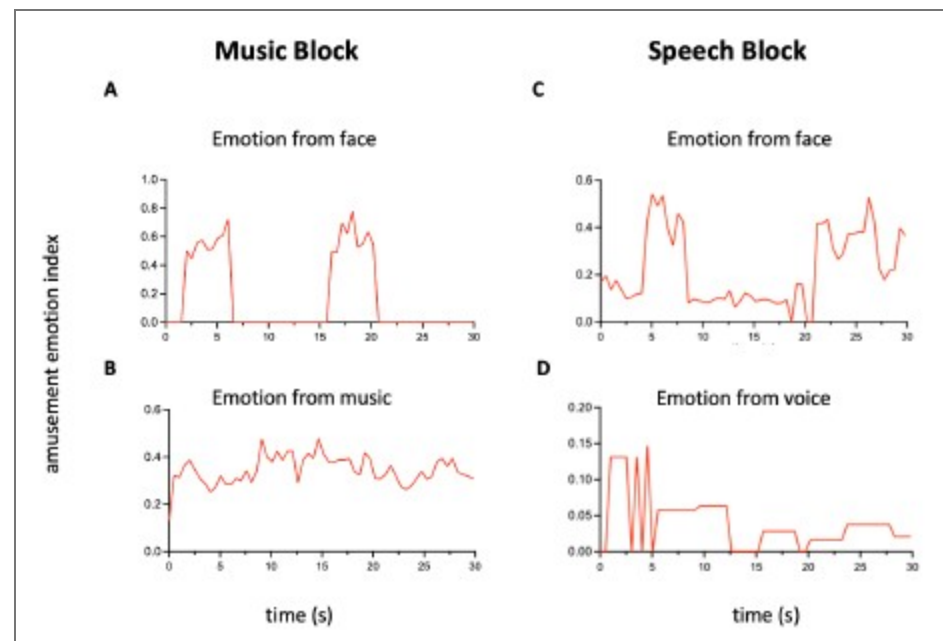
(1) The emotional stimuli included facial expressions embedded in speech or music, making it difficult to isolate neural responses to facial emotion per se from those related to speech content or music-induced emotion.

We thank the reviewer for their raising this important point. We agree that in naturalistic settings, face often co-occur with speech, and that these sources of emotion can overlap. However, background music induced emotions have distinct temporal dynamics which are separable from facial emotion (See the Author response image 2 (A) and (B) below). In addition, face can convey a wide range of emotions (48 categories in Hume AI model), whereas music conveys far fewer (13 categories reported by a recent study [1]). Thus, when using facial emotion feature time series as regressors (with 48 emotion categories and rapid

temporal dynamics), the model performance will reflect neural encoding of facial emotion in the music condition, rather than the slower and lower-dimensional emotion from music.

For the speech condition, we acknowledge that it is difficult to fully isolate neural responses to facial emotion from those to speech when the emotional content from faces and speech highly overlaps. However, in our study, (1) the time courses of emotion features from face and voice are still different (Author response image 2 (C) and (D)), (2) our main finding that DLPFC encodes facial expression information in postchildhood individuals but not in young children was found in both speech and music condition (Figure 2B and 2C). In music condition, neural responses to facial emotion are not affected by speech. Thus, we have included the DLPFC results from the music condition in the revised manuscript (Figure 2C), and we acknowledge that this issue should be carefully considered in future studies using videos with speech, as we have indicated in the future directions in the last paragraph of Discussion.

(1) Cowen, A. S., Fang, X., Sauter, D. & Keltner, D. What music makes us feel: At least 13 dimensions organize subjective experiences associated with music across different cultures. *Proc Natl Acad Sci USA* 117, 1924–1934 (2020).



Author response image 2. Time courses of the amusement. (A) and (B) Amusement conveyed by face or music in a 30-s music block. Facial emotion features are extracted by Hume AI. For emotion from music, we approximated the amusement time course using a weighted combination of low-level acoustic features (RMS energy, spectral centroid, MFCCs), which capture intensity, brightness, and timbre cues linked to amusement. Notice that music continues when there are no faces presented. (C) and (D) Amusement conveyed by face or voice in a 30-s speech block. From 0 to 5 seconds, a girl is introducing her friend to a stranger. The camera focuses on the friend, who appears nervous, while the girl's voice sounds cheerful. This mismatch explains why the shapes of the two time series differ at the beginning. Such situations occur frequently in naturalistic movies

(2) While the authors leveraged Hume AI to extract facial expression features from the video stimuli, they did not provide any validation of the tool's accuracy or reliability in the context of their dataset. It remains unclear how well the AI-derived emotion ratings align with human perception, particularly given the complexity and variability of naturalistic stimuli. Without such validation, it is difficult to assess the interpretability and robustness of the decoding results based on these features.

Hume AI models were trained and validated by human intensity ratings of large-scale, experimentally controlled emotional expression data [1-2]. The training process used both manual annotations from human raters and deep neural networks. Over 3000 human raters categorized facial expressions into emotion categories and rated on a 1-100 intensity scale. Thus, the outputs of Hume AI model reflect what typical facial expressions convey (based on how people actually interpret them), that is, the presented facial emotion. Our goal of the present study was to examine how facial emotions presented in the videos are encoded in the human brain at different developmental stages. We agree that the interpretation of facial emotions may be different in individual participants, resulting in different perceived emotion (i.e., the emotion that the observer subjectively interprets). Behavioral ratings are necessary to study the encoding of subjectively perceived emotion, which is a very interesting direction but beyond the scope of the present work. We have added text in the Discussion to explicitly note that our study focused on the encoding of presented emotion (second paragraph in Page 8).

(1) Brooks, J. A. et al. Deep learning reveals what facial expressions mean to people in different cultures. *iScience* 27, 109175 (2024).

(2) Brooks, J. A. et al. Deep learning reveals what vocal bursts express in different cultures. *Nat. Hum. Behav.* 7, 240–250 (2023).

(3) Only two children had relevant pSTC coverage, severely limiting the reliability and generalizability of results.

We appreciated this point and agreed with both reviewers who raised it as a significant concern. As described in response to reviewer 1 (comment 1), we have added data from another two children who have pSTC coverage. Group-level analysis using permutation test showed that children's pSTC significantly encode facial emotion in naturalistic contexts (Figure 3B). Because iEEG data from young children are extremely rare, rapidly increasing the sample size within a few years is not feasible. However, we are confident in the reliability of our conclusion that children's pSTC can encode facial emotion. First, the two new children's responses (S33 and S49) from pSTC were highly consistent with our previous observations (see individual data in Figure 3B). Second, the averaged prediction accuracy in children's pSTC ($r_{\text{speech}}=0.1565$) was highly comparable to that in post-childhood group ($r_{\text{speech}}=0.1515$).

(4) The rationale for focusing exclusively on high-frequency activity for decoding emotion representations is not provided, nor are results from other frequency bands explored.

We focused on high-frequency broadband (HFB) activity because it is widely considered to reflect the responses of local neuronal populations near the recording electrode, whereas low-frequency oscillations in the theta, alpha, and beta ranges are thought to serve as carrier frequencies for long-range communication across distributed networks[1-2]. Since our study aimed to examine the representation of facial emotion in localized cortical regions (DLPFC and pSTC), HFB activity provides the most direct measure of the relevant neural responses. We have added this rationale to the manuscript (Page 3).

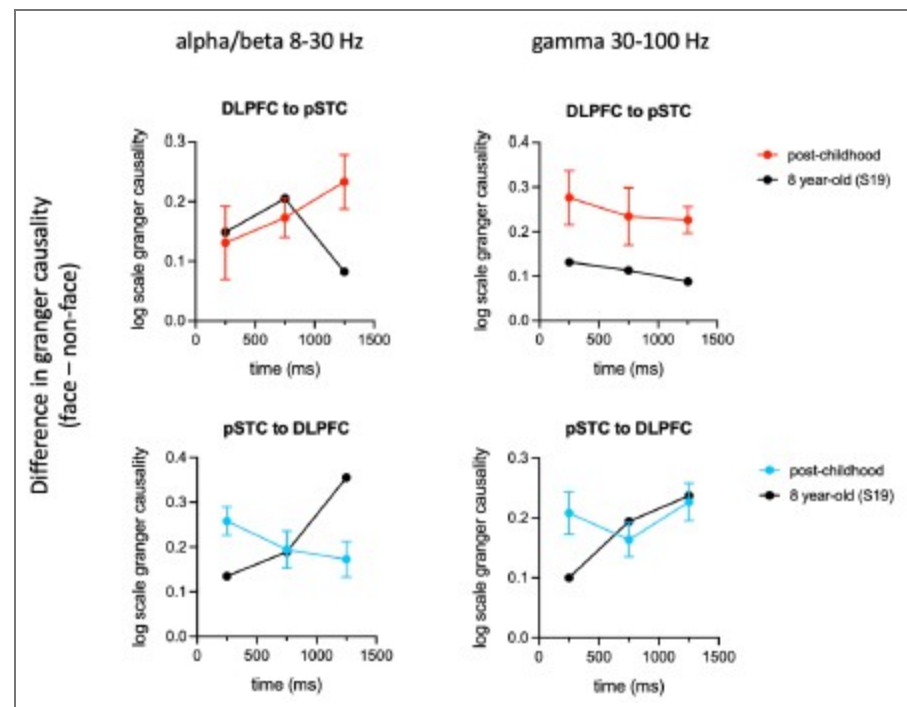
(1) Parvizi, J. & Kastner, S. Promises and limitations of human intracranial electroencephalography. *Nat. Neurosci.* 21, 474–483 (2018).

(2) Buzsaki, G. *Rhythms of the Brain*. (Oxford University Press, Oxford, 200ti).

(5) The hypothesis of developmental emergence of top-down prefrontal modulation is not directly tested. No connectivity or co-activation analyses are reported, and the number of participants with simultaneous coverage of pSTC and DLPFC is not specified.

Directional connectivity analysis results were not shown because only one child has simultaneous coverage of pSTC and DLPFC. However, the Granger Causality results from post-childhood group (N=7) clearly showed that the influence in the alpha/beta band from DLPFC to pSTC (top-down) is gradually increased above the onset of face presentation (Author response image 3, below left, plotted in red). By comparison, the influence in the alpha/beta band from pSTC to DLPFC (bottom-up) is gradually decreased after the onset of face presentation (Author response image 3, below left, blue curve). The influence in alpha/beta band from DLPFC to pSTC was significantly increased at 750 and 1250 ms after the face presentation (face vs nonface, paired t-test, Bonferroni corrected $P=0.005$, 0.006), suggesting an enhanced top-down modulation in the post-childhood group during watching emotional faces. Interestingly, this top-down influence appears very different in the 8-year-old child at 1250 ms after the face presentation (Author response image 3, below left, black curve).

As we cannot draw direct conclusions from the single-subject sample presented here, the top-down hypothesis is introduced only as a possible explanation for our current results. We have removed potentially misleading statements, and we plan to test this hypothesis directly using MEG in the future.



Author response image 3. Difference of Granger causality indices (face - nonface) in alpha/beta and gamma band for both directions. We identified a series of face onset in the movie that participant watched. Each trial was defined as -0.1 to 1.5 s relative to the onset. For the non-face control trials, we used houses, animals and scenes. Granger causality was calculated for 0-0.5 s, 0.5-1 s and 1-1.5 s time window. For the post-childhood group, GC indices were averaged across participants. Error bar is sem.

(6) The "post-childhood" group spans ages 13-55, conflating adolescence, young adulthood, and middle age. Developmental conclusions would benefit from finer age stratification.

We appreciate this insightful comment. Our current sample size does not allow such stratification. But we plan to address this important issue in future MEG studies with larger cohorts.

(7) The so-called "complex emotions" (e.g., embarrassment, pride, guilt, interest) used in the study often require contextual information, such as speech or narrative cues, for accurate interpretation, and are not typically discernible from facial expressions alone. As such, the observed age-related increase in neural encoding of these emotions may reflect not solely the maturation of facial emotion perception, but rather the development of integrative processing that combines facial, linguistic, and contextual cues. This raises the possibility that the reported effects are driven in part by language comprehension or broader social-cognitive integration, rather than by changes in facial expression processing per se.

We agree with this interpretation. Indeed, our results already show that speech influences the encoding of facial emotion in the DLPFC differently in the childhood and post-childhood groups (Figure 2D), suggesting that children's ability to integrate multiple cues is still developing. Future studies are needed to systematically examine how linguistic cues and prior experiences contribute to the understanding of complex emotions from faces, which we have added to our future directions section (last paragraph in Discussion, Page 8-9).

Recommendations for the authors:

Reviewer #1 (Recommendations for the authors):

In the introduction: "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." Aren't these supposed to be iEEG data instead of neuroimaging?

Corrected.

Reviewer #2 (Recommendations for the authors):

This manuscript would benefit from several improvements to strengthen the validity and interpretability of the findings:

(1) Increase the sample size, especially for children with pSTC coverage.

We added data from another two children who have pSTC coverage. Please see our response to reviewer 2's comment 3 and reviewer 1's comment 1.

(2) Include directional connectivity analyses to test the proposed top-down modulation from DLPFC to pSTC.

Thanks for the suggestion. Please see our response to reviewer 2's comment 5.

(3) Use controlled stimuli in an additional experiment to separate the effects of facial expression, speech, and music.

This is an excellent point. However, iEEG data collection from children is an exceptionally rare opportunity and typically requires many years, so we are unable to add a controlled-stimulus experiment to the current study. We plan to consider using controlled stimuli to study the processing of complex emotion using non-invasive method in the future. In addition, please see our response to reviewer 2's comment 1 for a description of how neural responses to facial expression and music are separated in our study.

<https://doi.org/10.7554/eLife.107636.2.sa0>