Behavioral/Cognitive

Development of Human Emotion Circuits Investigated Using a Big-Data Analytic Approach: Stability, Reliability, and Robustness

Yuan Zhang,[1] Aarthi Padmanabhan,[1] James J. Gross,[2] and Vinod Menon[1,3,4]

[1]Department of Psychiatry and Behavioral Sciences, [2]Department of Psychology, [3]Department of Neurology and Neurological Sciences, and [4]Stanford Neuroscience Institute, Stanford University, Stanford, California 94305

Emotion perception is fundamental to affective and cognitive development and is thought to involve distributed brain circuits. Efforts to chart neurodevelopmental changes in emotion have been severely hampered by narrowly focused approaches centered on activation of individual brain regions and small sample sizes. Here we investigate the maturation of human functional brain circuits associated with identification of fearful, angry, sad, happy, and neutral faces using a large sample of 759 children, adolescents, and adults (ages 8–23; female/male = 419/340). Network analysis of emotion-related brain circuits revealed three functional modules, encompassing lateral frontoparietal, medial prefrontal-posterior cingulate, and subcortical-posterior insular cortices, with hubs in medial prefrontal, but not posterior cingulate, cortex. This overall network architecture was stable by age 8, and it anchored maturation of circuits important for salience detection and cognitive control, as well as dissociable circuit patterns across distinct emotion categories. Our findings point to similarities and differences in functional circuits associated with identification of fearful, angry, sad, happy, and neutral faces, and reveal aspects of brain circuit organization underlying emotion perception that are stable over development as well as features that change with age. Reliability analyses demonstrated the robustness of our findings and highlighted the importance of large samples for probing functional brain circuit development. Our study emphasizes a need to focus beyond amygdala circuits and provides a robust neurodevelopmental template for investigating emotion perception and identification in psychopathology.

Key words: Big-Data; emotion perception and identification; functional networks; maturation; reliability; replicability

---

Significance Statement

Emotion perception is fundamental to cognitive and affective development. However, efforts to chart neurodevelopmental changes in emotion perception have been hampered by narrowly focused approaches centered on the amygdala and prefrontal cortex and small sample sizes. Using a large sample of 759 children, adolescents, and adults and a multipronged analytical strategy, we investigated the development of brain network organization underlying identification and categorization of fearful, happy, angry, sad, and neutral facial expressions. Results revealed a developmentally stable modular architecture that anchored robust age-related and emotion category-related changes in brain connectivity across multiple brain systems that extend far beyond amygdala circuits and provide a new template for investigation of emotion processing in the developing brain.

Introduction

Identification of distinct facial expressions, which is fundamental to recognizing the emotional state of others, begins in infancy and continues to develop throughout childhood and adolescence (Batty and Taylor, 2006; Thomas et al., 2007; Rodger et al., 2015; Leitzke and Pollak, 2016; Theurel et al., 2016). How emotions are identified and categorized in the developing brain remains a fundamental open question. Analysis of task-evoked brain circuits

Table 1. Studies that examined age-related changes in amygdala activity and connectivity underlying emotion perception[a]

Study	Sample size	Age (yr)	Contrast and no. of trials	Task	Findings
Gee et al. (2013)	45 (19F)	4–22	Fearful face versus fixation (no. of trials = 24)	Participants viewed fearful faces interspersed with neutral faces in one run, and viewed happy faces interspersed with neutral faces in the other run	Amygdala activity decreased with age; amygdala-vmPFC connectivity decreased with age
Kujawa et al. (2016)	61 (35F) TD 57 (34F) AD	7–25	Emotional faces (happy, angry, or fearful) versus shapes (no. of trials per emotion category = 12)	Participants were instructed to match emotion (happy, angry, or fearful) of the target face in face blocks and match shapes in shape blocks	No age-related effect in amygdala activity; amygdala-dACC connectivity decreased with age in TD but increased with age in AD for all emotions
Wu et al. (2016)	61 (35F)	7–25	Emotional faces (angry, fearful, or happy) versus shapes (no. of trials per emotion category = 12)	Participants were instructed to match emotion (angry, fearful, or happy) of the target face in face blocks and match shapes in shape blocks	No age-related effect in amygdala activity; amygdala-ACC/mPFC connectivity decreased with age for all emotions
Vink et al. (2014)	60 (28F)	10–24	Positive versus neutral (no. of trials = 32), negative versus neutral (no. of trials = 32)	Participants viewed each picture selected from the IAPS and label the valence (negative, neutral, or positive)	Amygdala activity decreased with age; amygdala-mOFC connectivity increased with age
Wolf and Herringa (2016)	24 (13F) TD 24 (16F) PTSD	8–18	Threat versus neutral (no. of trials = 16)	Participants viewed threat and neutral images selected from the IAPS and label the valence	No age-related effect in amygdala activity; amygdala-mOFC connectivity increased with age in TD but decreased with age in PTSD

[a]If not specified, healthy individuals were recruited in the listed studies. Functional connectivity was measured by PPI in Gee et al. (2013), Vink et al. (2014), and Wolf and Herringa (2016) and by gPPI in Kujawa et al. (2016) and Wu et al. (2016). AD, Anxiety disorder; IAPS, International Affective Picture System; PPI, psychophysiological interaction; PTSD, post-traumatic stress disorder; TD, typically developing. For MNI coordinates of amygdala and PFC regions used in these studies, see Table 4. For results of replication analysis, see Tables 5–7.

using a "Big-Data" approach has the potential to accelerate foundational knowledge of emotion perception circuit development and provides robust templates for understanding miswiring of functional circuits in developmental psychopathologies. Here, we leverage the Philadelphia Neurodevelopmental Cohort (PNC) (Satterthwaite et al., 2014) to address critical gaps in our knowledge regarding the functional circuits of emotion perception and their development.

Meta-analyses of neuroimaging studies in adults have identified widely distributed, and overlapping functional systems engaged by distinct emotion categories (Kober et al., 2008; Fusar-Poli et al., 2009). In particular, brain regions associated with three large-scale functional brain networks are relevant for emotion processing based on their extensive role in integrating salient cues across cortical and limbic systems. The "salience network" facilitates rapid detection of personally relevant and salient emotional cues in the environment (Seeley et al., 2007; Menon and Uddin, 2010; Menon, 2011), the "default mode network" is important for mentalizing and inferring the emotional states of others (Greicius et al., 2003; Blakemore, 2008), and the frontoparietal (FP) "central executive network" is involved in attention, reappraisal, and regulation of reactivity to emotional stimuli (Dosenbach et al., 2007; Gyurak et al., 2011; Tupak et al., 2014; Braunstein et al., 2017; Sperduti et al., 2017). However, previous neurodevelopmental studies of emotion perception have focused almost exclusively on the amygdala and its connectivity with the mPFC (Gee et al., 2013; Vink et al., 2014; Kujawa et al., 2016; Wolf and Herringa, 2016; Wu et al., 2016), with inconsistent findings (for a summary, see Table 1). While some studies found no age-related changes (Kujawa et al., 2016; Wolf and Herringa, 2016; Wu et al., 2016), others have reported increases (Vink et al., 2014; Wolf and Herringa, 2016) or decreases with age (Gee et al., 2013; Kujawa et al., 2016; Wu et al., 2016) in amygdala activation or connectivity with the mPFC. In contrast, intrinsic functional connectivity

studies have highlighted protracted developmental changes in distributed brain networks (Power et al., 2010; Menon, 2013), including developmental changes in amygdala connectivity with subcortical, paralimbic, and limbic structures, polymodal association, and ventromedial prefrontal cortex (vmPFC) (Qin et al., 2012). These findings highlight a critical need for characterizing the development of task-evoked functional brain circuits and network organization underlying emotion perception (Barrett and Satpute, 2013).
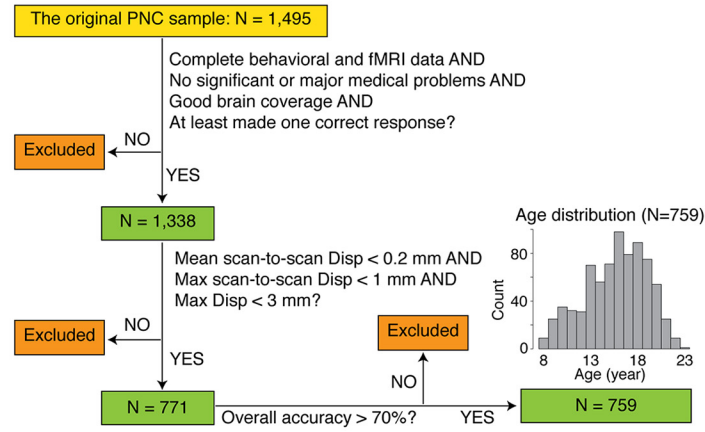
Here, we characterize the neurodevelopment of emotion perception circuits leveraging high-quality data ($N = 759$) from the emotion-identification fMRI task in the PNC, an unprecedented Big-Data initiative for charting brain development in a large community sample (Satterthwaite et al., 2014) (Fig. 1). We first used meta-analysis and community detection methods to identify and characterize the modular architecture of brain circuits associated with perception and identification of fearful, angry, sad, happy, and neutral faces. We then investigated aspects of brain circuit organization that are stable over development in addition to aspects that changed with age. Last, given the challenges of small sample sizes (Button et al., 2013; Szucs and Ioannidis, 2017; Turner et al., 2018), high false-positive rates (Eklund et al., 2016; Szucs and Ioannidis, 2017), and lack of replicability in neuroimaging research (Ioannidis, 2005; Eklund et al., 2016; Szucs and Ioannidis, 2017), we evaluated the reliability of our age-related findings on randomly subsampled independent datasets across a wide range of sample sizes (Fig. 1d) and assessed the replicability of previously reported findings. Our findings advance a new understanding of emotion circuit development in the human brain.
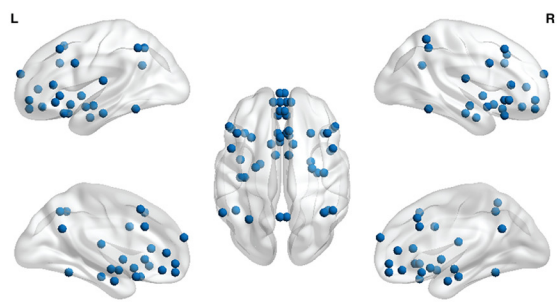
Materials and Methods

Participants
Participants were selected from the PNC 8–23 years of age ($N = 1495$). Data were excluded based on the following criteria: (1) lack of complete
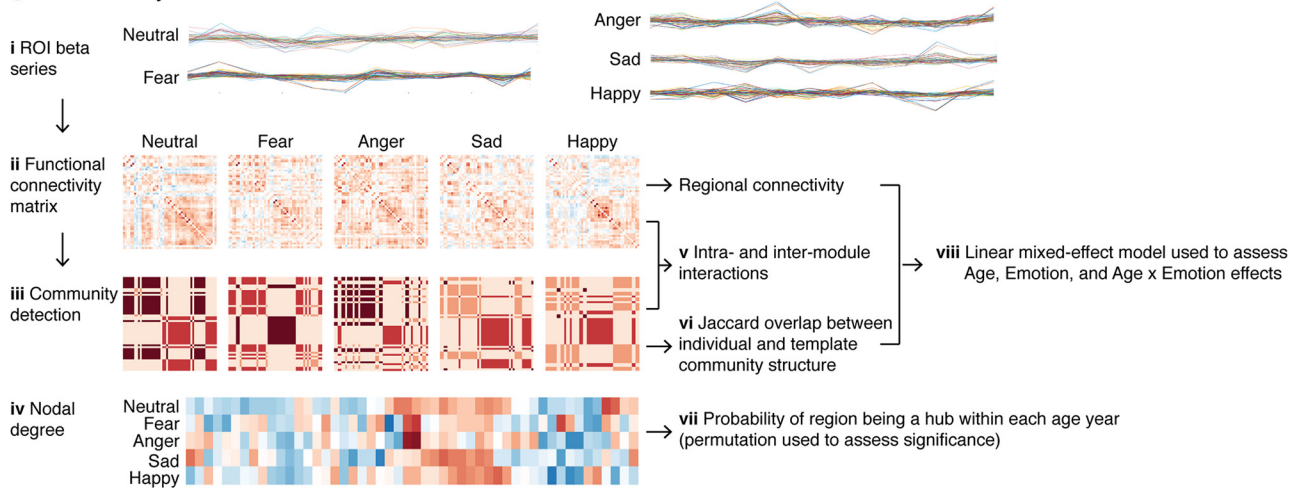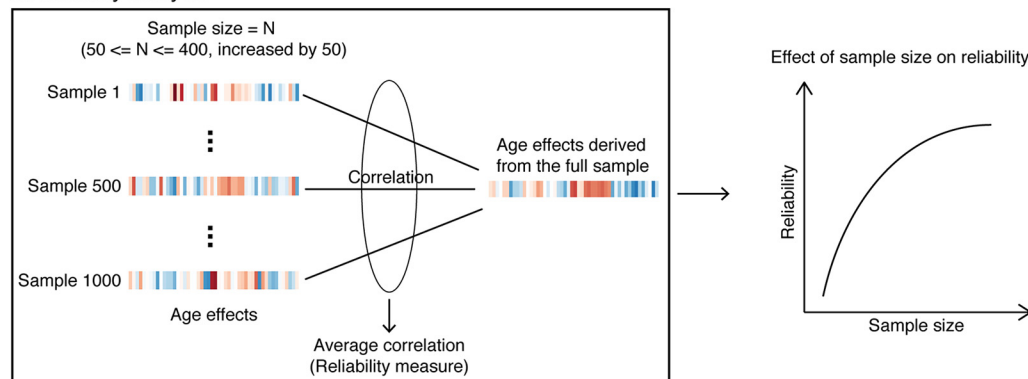
Figure 1. Schematic view of participant selection procedure and main analysis steps. ***a***, From the original 1495 participants, 157 were excluded due to incomplete data, medical problems, poor brain coverage, or not being able to perform the task. A total of 567 participants were excluded due to excess head motion, and 12 participants were excluded due to low task accuracy. A total of 759 participants 8 –23 years of age survived these exclusion criteria. ***b***, Brain ROIs involved in emotion perception were identified using meta-analysis implemented in Neurosynth (Yarkoni et al., 2011). They included the following: vmPFC, dmPFC, vlPFC, dlPFC, lOFC, IPL, SPL, sgACC, pgACC, dACC, PCC, pre-SMA, BLA, vAI, dAI, PI, hippocampus (Hipp), and FFG. Subcortical ROIs, including CMA, BNST, and NAc, were identified using anatomical atlases (see also Table 3). ROIs were overlaid on a reference brain surface using BrainNet Viewer (Xia et al., 2013). ***c***, Network analysis steps used to investigate development of the affective circuits. ***ci***, ROI β series were derived from each task condition (fear, anger, sad, happy, or neutral) for each individual. ***cii***, Functional connectivity matrices were computed for each individual. ***ciii***, Functional connectivity matrices were fed into a community detection algorithm to determine community structure of the ROIs. ***civ***, ROI nodal degree was computed over a wide range of sparsity (10% $\leq$ sparsity $\leq$ 40% with an interval of 5%), and then an integrated metric of nodal degree was obtained by computing the area under the curve. ***cv***, Intramodule and intermodule interactions were computed. ***cvi***, Deviation of individual community structure from template community structure was measured as the Jaccard overlap between the two. ***cvii***, ROIs were further categorized into hubs or nonhubs, and then the probability of an ROI to be a hub within each age year was calculated. Permutation test was used to assess hub significance. ***cviii***, These resulting brain metrics were further fed into linear mixed-effect models to assess age, emotion, and age $\times$ emotion effects. ***d***, Reliability analyses assessed the robustness of age-related findings and the effect of sample size. Briefly, we randomly drew N (sample size) participants from the full sample and assessed age-related changes in the subsample. This was repeated 1000 times across a wide range of sample sizes (50 $\leq N \leq$ 400, increased by 50). Correlation between age effects from a subsample and that from the full sample was calculated and then averaged across subsamples for each sample size. The resulting average correlation was used to measure reliability.

Table 2. Participant characteristics and task performance

Characteristic	Mean $\pm$ SD ($n = 759$)
Age (yr)	16.20 $\pm$ 3.27
Sex (female/male)	419/340
Race (Caucasian/other)	408/351
Motion (mm)	0.073 $\pm$ 0.03
Fluid intelligence[a]	12.62 $\pm$ 3.97
Fear accuracy	0.89 $\pm$ 0.11
Anger accuracy	0.92 $\pm$ 0.09
Sad accuracy	0.88 $\pm$ 0.11
Neutral accuracy	0.91 $\pm$ 0.09
Happy accuracy	0.99 $\pm$ 0.03
Fear RT (s)	2.65 $\pm$ 0.46
Anger RT (s)	2.48 $\pm$ 0.44
Sad RT (s)	2.68 $\pm$ 0.46
Happy RT (s)	2.09 $\pm$ 0.40
Neutral RT (s)	2.34 $\pm$ 0.44

[a]Fluid intelligence is measured by Penn Matrix Reasoning Test scores (i.e., total correct responses for all test trials). RT, Response time (median). For age distribution, see also Figure 1a.

emotion task fMRI data; (2) high in-scanner motion (mean scan-to-scan displacement > 0.2 mm or maximum scan-to-scan displace > 1 mm or maximum displacement > 3 mm); (3) missing brain coverage; (4) medical rating > 2 (i.e., having significant or major medical problems); and (5) overall task performance (accuracy) < 70%. A total of 759 participants (fMRI sample) survived these exclusion criteria and were included in further behavioral and neuroimaging analyses (Fig. 1a; Table 2). Additional behavioral analysis was performed to investigate age-related effects on emotion categorization accuracy and reaction time (RT) in a larger sample of 1338 participants (behavioral sample) who had valid behavioral data. A total of 157 participants with medical issues or incomplete data or were unable to perform the task were excluded from any analysis. All study procedures were approved by the institutional review boards of both the University of Pennsylvania and the Children's Hospital of Philadelphia (Satterthwaite et al., 2014).

Behavioral and demographic differences between the neuroimaging and excluded samples

We examined whether participants in the fMRI ($N = 759$) and excluded ($N = 579$) samples are similar cross-sections of the overall population in terms of age, sex, and behavioral performance (accuracy and RT). The excluded sample included participants who did not meet our criteria for inclusion in fMRI data analysis but were included in the behavioral analysis.

PNC emotion identification fMRI task

Participants were presented with faces and were required to categorize them as one of five emotion categories: neutral, fear, anger, sadness, or happiness. Each face was presented for 5.5 s followed by a jittered interstimulus interval ranging between 0.5 and 18.5 s, during which a complex crosshair that matched perceptual qualities of the face was displayed. There were 12 faces per emotion category, resulting in 60 total trials (Satterthwaite et al., 2014).

Neuroimaging data acquisition and preprocessing

Images were acquired using a 3T Siemens TIM Trio scanner (32 channel head coil), with a single-shot, interleaved multislice, gradient-echo, EPI sequence (TR/TE = 3000/32 ms, flip angle = 90°, FOV = 192 × 192 mm^2, matrix = 64 × 64, slice thickness = 3 mm). fMRI data were preprocessed and analyzed using Statistical Parametric Mapping software SPM8 (Wellcome Trust Department of Cognitive Neurology, London). Images were slice-time corrected, motion-corrected, resampled to 3 mm × 3 mm × 3 mm isotropic voxels, normalized to MNI space, spatially smoothed with a 6 mm FWHM Gaussian filter, and temporally filtered using a high-pass filter with a cutoff frequency of 1/128 Hz.

Head motion

Participants with mean scan-to-scan displacement ≤0.2 mm and maximum scan-to-scan displace ≤1 mm and maximum displacement ≤3

mm were included in further analysis. In addition to correction for motion correction at the whole-brain level during fMRI preprocessing (as described above), to rule out any potential residual effects of motion, mean scan-to-scan displacement was included as a nuisance covariate in all our analyses.

Network analysis

Network nodes. To overcome limitations of extant neurodevelopmental studies of emotion perception and categorization, which have predominantly focused on the amygdala (Gee et al., 2013; Vink et al., 2014; Kujawa et al., 2016; Wolf and Herringa, 2016; Wu et al., 2016), we conducted a meta-analysis to identify cortical and subcortical brain regions that are consistently involved in emotion-related processes. Specifically, using the meta-analysis toolbox Neurosynth (Yarkoni et al., 2011) and the search term "emotion," we identified local peaks (at least 8 mm apart) within each cluster in the forward and reverse inference maps. We selected peaks in frontal, parietal, medial temporal, limbic, and subcortical areas. ROIs were defined as a 6 mm sphere with an identified peak as center. If a unilateral peak was detected at a specific region, we included its contralateral counterpart by flipping the peak coordinate along the x axis.

Together, nodes encompassed the vmPFC, dorsomedial PFC (dmPFC), ventrolateral PFC (vlPFC), and dorsolateral PFC (dlPFC), lateral orbitofrontal cortex (lOFC), inferior parietal lobule (IPL), and superior parietal lobule (SPL), subgenual ACC (sgACC), pregenual ACC (pgACC), dorsal ACC (dACC), and posterior cingulate cortex (PCC), presupplementary motor area (pre-SMA), BLA, dorsal anterior insula (dAI), posterior insula (PI), hippocampus (Hipp), and fusiform gyrus (FFG). An additional pair of dorsolateral prefrontal ROIs superior to the pair identified by Neurosynth were included based on two recent emotion-regulation meta-analyses (Buhle et al., 2014; Kohn et al., 2014). We also included nodes in the ventral anterior insula (vAI), centromedial amygdala (CMA), bed nucleus stria terminalis (BNST), and nucleus accumbens (NAc) based on their well-known roles in perception of negative and positive emotions (LeDoux, 2007; Deen et al., 2011; Chang et al., 2013; Floresco, 2015; Lebow and Chen, 2016; Namkung et al., 2017). The vAI node was identified using Neurosynth and search term "anxiety" (Paulus and Stein, 2006; Etkin, 2010; Menon, 2011; Blackford and Pine, 2012; Qin et al., 2014), the CMA was identified based on its cytoarchitectonically defined probabilistic map using the Anatomy toolbox (Eickhoff et al., 2005), the BNST node was based on the meta-analysis by Avery et al. (2014), and NAc was identified based on the Harvard-Oxford atlas distributed with FSL (http://www.fmrib.ox.ac.uk/fsl/). These procedures identified 50 ROIs consistently associated with emotion processing (Fig. 1b; Table 3). Finally, we further verified that the activation maps of emotion-related topics encompassed the maps reported in a large meta-analytic database of fMRI studies (Rubin et al., 2017).

Emotion-related node-level activity. Task-related brain activation was identified using a GLM implemented in SPM8. At individual level, brain activation representing correct trials for each emotion task condition (neutral, fear, anger, sad, or happy) was modeled using boxcar functions convolved with a canonical hemodynamic response function and a temporal derivative to account for voxelwise latency differences in hemodynamic response. An error regressor was also included in the model to account for the influence of incorrect trials. Additionally, six head movement parameters generated from the realignment procedure were included as regressors of no interest. Serial correlations were accounted for by modeling the fMRI time series as a first-degree autoregressive process. For each emotion task condition, ROI activity was obtained by averaging the estimated β values across voxels within that ROI.

Emotion-related internode connectivity. Functional connectivity was assessed using a β series analysis approach (Rissman et al., 2004). At individual level, brain activation related to each trial was modeled using boxcar functions convolved with a canonical hemodynamic response function. Six head movement parameters were included as regressors of no interest. Serial correlations were accounted for by modeling the fMRI time series as a first-degree autoregressive process. Trialwise ROI activity was obtained by averaging the estimated β values across voxels within that ROI. A β series was derived for each ROI each emotion condition by concatenating trialwise ROI activity belonging to the same emotion con-

Table 3. Anatomical location and MNI coordinates of emotion circuit-related nodes

Region/abbreviation	MNI coordinates			Brodmann area
	x	y	z	
Left BLA (BLA.L)	−20	−6	−20	
Right BLA (BLA.R)	26	−4	−20	
Left CMA (CMA.L)	−24	−9	9	
Right CMA (CMA.R)	27	−10	−9	
Left sgACC (sgACC.L)	−2	14	−16	BA 25
Right sgACC (sgACC.R)	6	20	−12	BA 25
Left pgACC (pgACC.L)	−2	42	6	BA 32
Right pgACC (pgACC.R)	2	42	8	BA 32
Left dACC (dACC.L)	−4	22	30	BA 32
Right dACC (dACC.R)	4	22	30	BA 32
Left PCC (PCC.L)	−2	−54	28	BA 31
Right PCC (PCC.R)	4	−54	28	BA 31
Left vmPFC (vmPFC.L)	−8	50	−4	BA 10
Right vmPFC (vmPFC.R)	8	50	−4	BA 10
Left vmPFC (vmPFC.L)	−2	50	−12	BA 32
Right vmPFC (vmPFC.R)	4	50	−12	BA 32
Left vmPFC (vmPFC.L)	−4	38	−10	BA 32
Right vmPFC (vmPFC.R)	4	38	−10	BA 32
Left dmPFC (dmPFC.L)	−2	58	20	BA 9
Right dmPFC (dmPFC.R)	2	58	20	BA 9
Left pre-SMA (preSMA.L)	−2	18	46	BA 6
Right pre-SMA (preSMA.R)	6	24	38	BA 32
Left lOFC (lOFC.L)	−46	26	−6	BA 47
Right lOFC (lOFC.R)	46	26	−4	BA 47
Left lOFC (lOFC.L)	−28	18	−16	BA 47
Right lOFC (lOFC.R)	26	22	−18	BA 47
Left vlPFC (vlPFC.L)	−48	28	10	BA 45
Right vlPFC (vlPFC.R)	48	28	10	BA 45
Left dlPFC (dlPFC.L)	−46	8	30	BA 9
Right dlPFC (dlPFC.R)	46	8	30	BA 9
Left dlPFC (dlPFC.L)	−41	22	43	BA 8
Right dlPFC (dlPFC.R)	41	22	43	BA 8
Left IPL (IPL.L)	−54	−50	44	BA 40
Right IPL (IPL.R)	46	−48	44	BA 40
Left SPL (SPL.L)	−30	−56	44	BA 7
Right SPL (SPL.R)	38	−50	52	BA 7
Left FFG (FFG.L)	−42	−48	−20	BA 37
Right FFG (FFG.R)	42	−48	−20	BA 37
Left dAI (dAI.L)	−34	22	0	BA 13
Right dAI (dAI.R)	38	22	−4	BA 13
Left vAI (vAI.L)	−38	−2	−10	BA 13
Right vAI (vAI.R)	44	6	−10	BA 13
Left PI (PI.L)	−38	−18	14	BA 13
Right PI (PI.R)	38	−14	16	BA 13
Left BNST (BNST.L)	−6	2	0	
Right BNST (BNST.R)	6	2	0	
Left NAc (NAc.L)	−10	12	−8	
Right NAc (NAc.R)	10	12	−8	
Left hippocampus (Hipp.L)	−34	−18	−16	
Right hippocampus (Hipp.R)	32	−14	−16	

dition. Only ROI activity corresponding to correct trials was used to construct β series. Conditionwise β series were demeaned and used to calculate Pearson correlation between all pairs of ROIs. Pearson correlation coefficients were Fisher Z-transformed and used to index functional connectivity strength and form emotion-related networks in each participant (Fig. 1c).

Emotion-related network structure. To determine common patterns of community structure across all participants, we used a consensus-based approach for identifying functional modules (Lancichinetti and Fortunato, 2012; Dwyer et al., 2014) associated with each emotion. In each participant, the Louvain algorithm (Blondel et al., 2008), as implemented in the Brain Connectivity Toolbox (Rubinov and Sporns, 2010), was used to first determine the modular structure of the signed, weighted connectivity matrix (Newman, 2004, 2006). To handle potential degeneracy of community assignments, we repeated this procedure 1000 times in each

participant (Lancichinetti and Fortunato, 2012; Dwyer et al., 2014). In each iteration, we generated a co-classification matrix in which each element contained 1 if two nodes were part of the same module and 0 otherwise. The resulting 1000 co-classification matrices were averaged to generate a co-occurrence matrix indicating the probability of two nodes being in the same module across 1000 iterations. Co-occurrence matrices were then averaged across participants to produce a positive, weighted group consistency matrix, with higher weight in elements of this matrix indicating that two nodes were more frequently classified in the same module across participants. This group consistency matrix was used to determine a consensus partition (Lancichinetti and Fortunato, 2012). To evaluate the influence of resolution parameter gamma, which tunes the size and number of communities obtained from modularity maximization (Fortunato and Barthélemy, 2007; Sporns and Betzel, 2016), the above consensus-based approach was repeated at gamma values ranging from 0.1 to 2.0 in increments of 0.1.

To assess the significance of the modular structure, we compared the actual Q to the Q derived from null networks. Specifically, the group consistency matrix was used to create 1000 null networks with preserved weight and degree distributions. For each null network, weak edges with weight <0.1 were first set to 0 and Q was calculated 1000 times and used to establish an empirical null distribution of Q. Q was considered significant if it was larger than the 95th percentile of the empirical null distribution.

Emotion-related hubs. To identify hubs in the affective circuits and their development with age, we first computed nodal degree for each individual and each emotion category. Degree of a node i is defined as the number of edges connected to that node as follows:

$$k_i = \sum_{j \in n, \, j \neq i} e_{ij}$$

where e_{ij} is the connection status (0 or 1) between node i and node j, and n is the total number of nodes (Rubinov and Sporns, 2010). Because the functional matrix is fully connected, a sparsity condition, based on the ratio of the actual edge number to the maximum possible edge number in a network, was applied to the connectivity matrix before computing nodal degree. To minimize the impact of setting an arbitrary level of sparsity, we calculated degree over a wide range of sparsity (i.e., 10% ≤ sparsity ≤ 40% with an interval of 5%) and then computed area under the curve, to obtain an integrated metric of node degree as in previous studies examining functional circuits (Yan et al., 2013; Lv et al., 2016). A node with high degree indicates high interconnectivity with other brain nodes. Hub nodes were defined as nodes with integrated degree metric at least 1 SD above the mean across all nodes.

Generalization of ROIs to the Brainnetome atlas. We used the Brainnetome atlas to determine robustness and generalizability of findings from the 50 emotion-related meta-analysis-derived ROIs. An advantage of the Brainnetome atlas, compared with many other functional brain parcellations, is that it includes both cortical and subcortical clusters with appropriate anatomical labels (Fan et al., 2016). Specifically, we identified Brainnetome clusters in which each of the 50 ROIs was located; because 2 ROI pairs were located in the same Brainnetome cluster, this procedure yielded 46 matching clusters. We then repeated all our analyses using these spatially extended clusters.

Statistical analysis
Linear versus nonlinear age models. The remainder of our analyses involved assessing age-related changes in task performance, regional activity, regional connectivity, and intramodule and intermodule connectivity. For task performance, we analyzed both the fMRI ($N = 759$) and behavioral ($N = 1338$) samples. We examined three models (linear, quadratic, and cubic spline) to test for linear and nonlinear age-related effects, and used Bayesian Information Criteria to determine the best model. The vast majority of the analyses indicated that nonlinear models did not fit the data better than linear models. Where appropriate, we report nonlinear age effects that survived FDR correction for multiple comparisons.

Linear mixed-effect model. To examine main effects of age, emotion, and age × emotion interactions on task performance and brain measures, we used R package lme4 (Bates et al., 2015) to perform a linear

mixed-effect analysis. As fixed effects, we entered age, emotion, and age × emotion interactions (regressors of interest) as well as sex, race, and head movement (regressors of no interest) into the model. As random effects, we had intercepts for participants. Significance of regressor coefficients were obtained by F tests.

Multiple linear regression of age-related changes. For significant age × emotion interactions, we performed additional analyses to examine age-related changes in each emotion category. Specifically, we used multiple linear regression models, with behavioral or brain measure as dependent variable, age as a covariate of interest, and sex, race, and head movement as covariates of no interest.

Development of emotion circuits

Development of emotion-related network structure. To determine how network structure changes with age, we first determined community structure for each participant and each group as described above. We then computed the Jaccard index between individual and group consensus community structures. Specifically, overlap in community structure was computed using a co-occurrence matrix, with 1 indicating two nodes belonging to the same module and 0 indicating two nodes belonging to different modules. The Jaccard index is defined as the ratio of the intersection between two matrices divided by their union, ranging from 0 (no overlap) to 1 (full overlap). A linear mixed-effect model was then applied to the Jaccard index to determine main effects of age, emotion, and age × emotion interactions. We also examined community structure in three distinct age groups, children (8–12 years), adolescents (13–17 years), and adults (18–23 years), to provide additional insights into the development of emotion circuits.

Development of emotion-related intramodule and intermodule connectivity. To examine main effect of age, emotion, and age × emotion interactions on functional interactions between the module, we assessed the intramodule connectivity, by computing the mean strength of all connectional weights within a module, and intermodule interaction, by computing the mean strength of connectional weights between modules, for each individual and each emotion category.

Development of emotion-related hubs. To investigate the development of hub nodes, we first computed the probability of an ROI being a hub in a 1 year age window, separately for each emotion category. Specifically, for each participant and each emotion category, we identified hubs (see above) and counted, for each node, how many times it functioned as hub across participants under a specific emotion category. This count was divided by the total number of participants within that specific age window to derive a node's hub probability. To assess the significance of hub probability, we constructed an empirical null distribution for each region and each age window under each emotion category. Specifically, we first permuted hub identity across all nodes within each individual, by randomly designating which nodes were hubs while keeping the total number of hubs constant over 1000 iterations, and then calculated the probability of a brain region being a hub node within each age window for all permutations. Only nodes with a probability above the 95th percentile of its empirical null distribution were considered as a significant hub.

Classification analysis

We used a nonlinear support vector machine classification algorithm with 10-fold cross-validation to assess the discriminability of regional connectivity measures between all pairs of emotion categories. The python scikit-learn package (https://scikit-learn.org/stable/) was used to perform this analysis. Permutation test was used to assess the performance of the classifier. Specifically, for each pair of emotion categories, data were divided into 10 folds, with those belonging to the same participant included in the same fold. A nonlinear support vector machine classifier was trained using nine folds, leaving out one fold. The samples in the left-out fold were then predicted using this classifier, and an accuracy was derived by calculating the percentage of correct predictions. This procedure was repeated 10 times, and a final classification accuracy was derived by calculating the mean of the 10 left-out accuracy, which was used as a measure of model performance. Next, the statistical significance of the model was assessed using permutation test. In each permu-

tation, emotion category labels (e.g., neutral vs fear) were randomly exchanged with a probability of 0.5 for each participant. A classifier was trained, and a classification accuracy was derived using the same scheme described above. Classification accuracy from 100 permutations was used to construct the empirical null distribution from which a p value was computed.

Reliability analysis

Behavior. We examined the reliability of our developmental findings and assessed the effect of sample size on reliability for each emotion category. First, we randomly drew N (sample size) participants ($50 \leq N \leq 700$, incremented by 50) from either the fMRI ($N = 759$), or behavioral ($N = 1338$), sample without replacement and with the proportion of children (ages 8–12 years), adolescents (ages 13–17 years), and adults (ages 18–23 years) in each subsample being matched to the fMRI, or behavioral, sample. Second, we assessed developmental changes in the subsample using multiple linear regression models. Third, we repeated the first two steps 1000 times at each sample size. Finally, we calculated the probability of observing a significant age effect with the same direction as the one revealed in the full sample across 1000 runs for each sample size, which was used as the measure of reliability. We defined findings with reliability >0.7 as reliable because it is unlikely ($p < 2.2e-16$) to observe a significant effect ≥ 700 times of 1000 runs, assuming equal probability to observe a significant or null effect in a run.

Brain networks. We randomly drew N participants from the fMRI sample ($N = 759$) and assessed age-related changes in the subsample using linear mixed-effect model. This was repeated 1000 times at each sample size across a wide range of sample sizes ($50 \leq N \leq 400$). Next, correlation between age effects in regional activity or connectivity across all brain regions from a subsample and that from the full sample was calculated and then averaged across subsamples at each sample size. The resulting average correlation was used as the measure of network-level reliability of overall developmental pattern across all brain regions. We defined reliable findings as those with correlation >0.7 (i.e., $R^2 > 0.5$), the point at which the model explains more variance than is unexplained. Finally, we also constructed empirical null distributions for reliability measured as average correlation, using the same approach described above, with an additional step that age was permuted in each subsample before assessing developmental changes. Average correlation above the 95th percentile of its empirical null distribution was considered as significantly different from null.

Replication analysis of previous developmental findings related to amygdala circuitry

We examined the development of amygdala activity and connectivity in the current PNC cohort using coordinates reported in three published task-fMRI studies using similar emotion face perception tasks, contrasts, and age range as the PNC (Gee et al., 2013; Kujawa et al., 2016; Wu et al., 2016). Tables 1 and 4 provide details of the studies and coordinates used. Moreover, we performed additional generalized psychophysiological interaction (gPPI) analysis to establish amygdala-mPFC connectivity and examine age-related changes as a similar method (gPPI or PPI) was used in the three published studies. Specifically, the gPPI model included one regressor of seed time series, five regressors of task conditions (neutral, fear, anger, sadness, or happiness), and their interaction with seed time series, one regressor of error trials and its interaction with seed time series, six regressors of head movement, and one regressor of constant. Replication analysis was then applied to amygdala-mPFC gPPI connectivity.

Data and code availability

Raw data associated with this work are available through dbGaP (https://www.ncbi.nlm.nih.gov/projects/gap/cgi-bin/study.cgi?study_id=phs000607.v2.p2).

All data analysis scripts used in the study will be made publicly available upon publication.

Results

Behavioral maturation of emotion identification

We first examined the main effects of age, emotion, and age × emotion interactions on accuracy (i.e., proportion of correct re-

Table 4. MNI coordinates of amygdala and PFC regions used in previous developmental studies of face emotion perception[a]

Study	Amygdala (x, y, z)	Connectivity Seed (x, y, z) and target (x, y, z)
Gee et al. (2013)	$(32, -1, -16)$	Right amygdala $(32, -1, -16)$ and right vmPFC $(2, 32, 8)$
Kujawa et al. (2016)	$(-20, -2, -20)$, $(24, -2, -22)$	Left amygdala (AAL) and left dACC $(-4, 30, 16)$
		Right amygdala (AAL) and right dACC $(2, 34, 14)$
Wu et al. (2016)	Left and right amygdala from AAL atlas	Left amygdala (AAL) and left ACC/mPFC $(-6, 34, 16)$
		Right amygdala (AAL) and left ACC/mPFC $(-4, 36, 14)$
		Left amygdala (AAL) and left ACC $(-8, 28, 18)$
		Right amygdala (AAL) and right ACC $(6, 36, 12)$

[a]AAL, Automated anatomical labeling. While no age-related effects were revealed in amygdala activity in Kujawa et al. (2016) and Wu et al. (2016), we created amygdala ROI based on coordinates showing significant main effect of emotion category in Kujawa et al. (2016) or used amygdala AAL atlas, which was a seed region in connectivity analysis in Wu et al. (2016), to examine age-related effects of amygdala activity in our PNC cohort.

sponses) and median RT of emotion identification in the fMRI sample ($N = 759$), which included participants who had valid behavioral and neuroimaging data.

Accuracy versus age
Our analysis revealed no main effect of age ($p > 0.1$) but a significant effect of emotion ($F_{(4,3028)} = 187.156$, $p < 2.2e-16$) and age × emotion interactions ($F_{(4,3028)} = 5.370$, $p = 0.0003$). Further analysis revealed that accuracy increased with age for fear ($r = 0.151$, $p = 4.71e-05$) and anger ($r = 0.152$, $p = 3.83e-05$) conditions (Fig. 2ai), but not for sad ($r = -0.028$, $p = 0.457$), happy ($r = 0.064$, $p = 0.084$), or neutral ($r = 0.055$, $p = 0.144$) conditions.

RT versus age
Results revealed a significant age effect ($F_{(1,1785)} = 5.370$, $p = 1.15e-07$), showing that median RT decreased with age such that older participants were faster at identifying emotions. Moreover, there is a significant effect of emotion ($F(4, 3028) = 397.167$, $p < 2.2e-16$) and age × emotion interactions ($F_{(4,3028)} = 4.023$, $p = 0.003$). Further analysis revealed that median RT decreased with age for all five emotions (Fig. 2aii): fear ($r = -0.230$, $p = 4.62e-10$), anger ($r = -0.173$, $p = 3.11e-6$), sad ($r = -0.111$, $p = 0.003$), happy ($r = -0.235$, $p = 1.75e-10$), and neutral ($r = -0.207$, $p = 2.01e-08$). The significant age × emotion interaction was characterized by differences in slope between the sad and fear conditions and between the sad and happy conditions.

Reliability of age-related changes
We examined the reliability of age-related changes in relation to sample size. For each emotion category, we used multiple linear regression model to assess age-accuracy and age-RT associations across a wide range of sample sizes ($50 \leq N \leq 700$, increased by 50) with the proportion of children (ages 8–12 years), adolescents (ages 13–17 years), and adults (ages 18–23 years) in each subsample matched to the full sample (1000 times per sample size). Reliability was measured as probability of observing a significant correlation between age and accuracy or RT with the same direction as the one revealed in the full sample. For fear and anger, the two emotion categories that showed a significant age-related increase in accuracy in the full sample, reliability increased with sample size (Fig. 2ci). Reliability of negative age-RT

associations increased with sample size across all five emotion categories (Fig. 2cii). In general, a sample size of 300–350 was needed to identify reliable age-accuracy associations, and a sample size of 200–300 was needed to identify reliable age-RT association. Results also demonstrate that age-related improvements in emotion identification were strongest for fear and anger, the two high arousal-negative emotions.

Behavioral maturation of emotion identification assessed using the full behavioral sample
Most developmental neuroimaging studies typically restrict analysis of fMRI-task behavioral data to participants with high levels of accuracy and low head movement during scanning, as noted above. The large PNC dataset allowed us to investigate potential biases in age-related effects in a robust manner using a larger sample of 1338 participants.

Accuracy versus age
Results revealed significant main effect of age ($F_{(1,4515)} = 53.814$, $p < 2.6e-13$), showing that accuracy increased with age such that older participants were better at identifying emotions. Moreover, there was a significant main effect of emotion ($F_{(4,5344)} = 344.866$, $p < 1.86e-264$) and age × emotion interactions ($F_{(4,5344)} = 10.434$, $p < 2.04e-08$). Further analysis revealed that accuracy increased with age for neutral ($r = 0.078$, $p = 0.006$), fear ($r = 0.150$, $p = 9.32–08$), anger ($r = 0.129$, $p = 6.43e-06$), and happy ($r = 0.058$, $p = 0.044$) conditions, but not for the sad condition ($r = 0.036$, $p = 0.203$) (Fig. 2bi). Additional analysis revealed that the age × emotion interaction was driven by slope differences between fear and each of the other four conditions, between anger and happy conditions, and between neutral and happy conditions. While all significant findings revealed in the fMRI sample were replicated in the larger behavioral sample, we also noted a few additional significant effects in the behavioral sample, including a main effect of age as well as associations between accuracy and age under neutral and happy conditions. However, these additional age-related effects during neutral and happy conditions were weaker than those during fear and anger.

RT versus age
Results revealed a significant main age effect ($F_{(1,3091)} = 139.818$, $p = 1.38e-31$), showing that median RT decreased with age such that older participants were faster at identifying emotions. Moreover, there was a significant main effect of emotion ($F_{(4,5344)} = 594.997$, $p = 0.000$) and age × emotion interactions ($F_{(4,5344)} = 9.162$, $p = 2.25e-07$). Further analysis revealed that median RT decreased with age for all five emotions (Fig. 2bii): neutral ($r = -0.300$, $p = 9.57e-28$), fear ($r = -0.146$, $p = 1.85e-21$), anger ($r = -0.221$, $p = 4.76e-15$), sad ($r = -0.185$, $p = 4.84e-11$), and happy ($r = -0.260$, $p = 5.66e-21$). Additional analysis revealed that age × emotion interaction was driven by slope differences between fear and anger/sad/happy conditions and between neutral and anger/sad/happy conditions. Notably, all significant findings revealed in 759 participants were replicated in this larger behavioral sample.

Reliability of age-related changes
Reliability of age-related changes as a function of sample size were conducted using the same procedures as ones used above for the fMRI sample. For fear and anger, the two emotion categories that showed a significant age-related increase in accuracy in both the fMRI and the full behavioral samples, reliability increased with sample size; and in general, a sample size of 300–350 was needed to identify reliable age-accuracy associations (Fig. 2ci). This find-
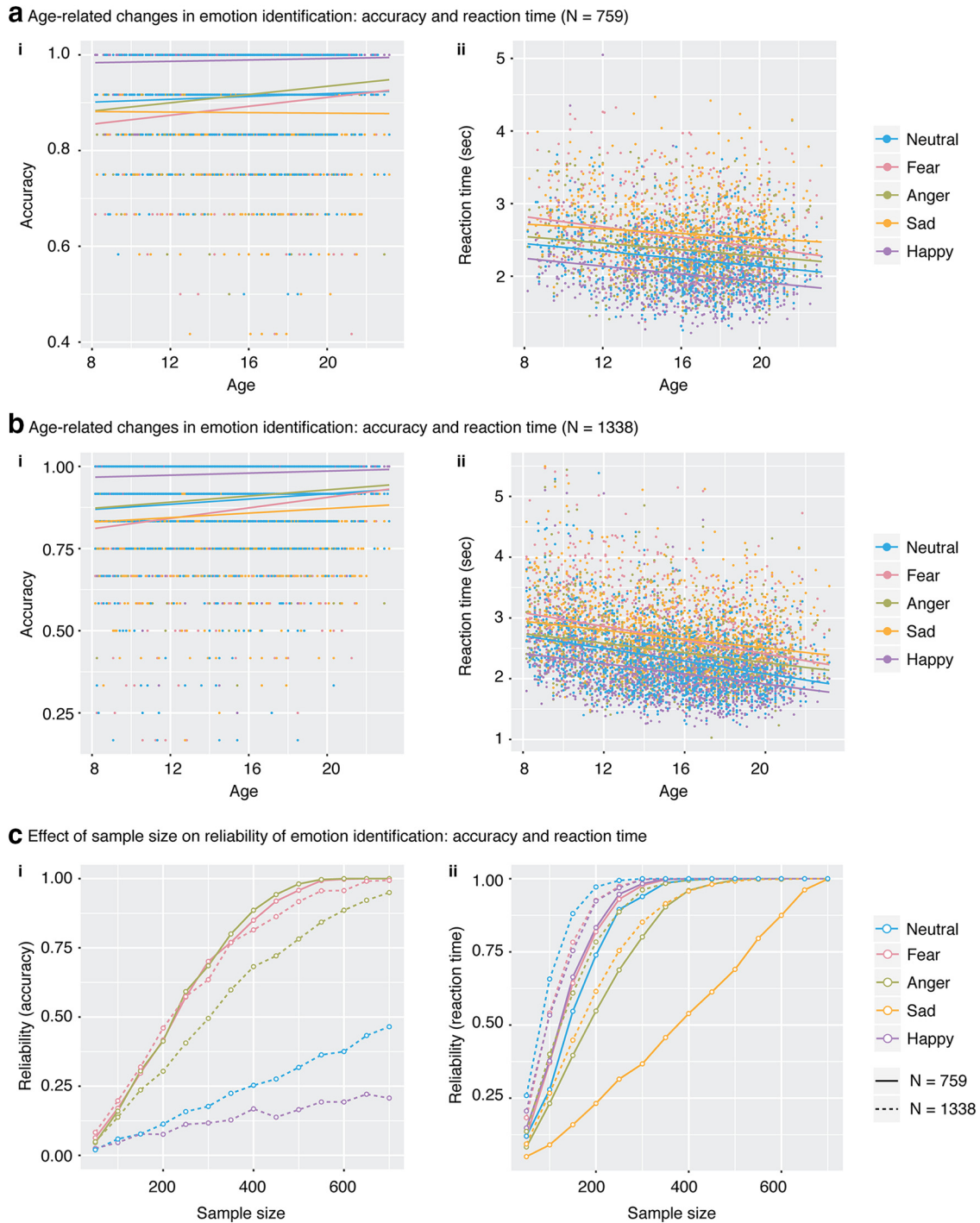
a Age-related changes in emotion identification: accuracy and reaction time (N = 759)

b Age-related changes in emotion identification: accuracy and reaction time (N = 1338)

c Effect of sample size on reliability of emotion identification: accuracy and reaction time

Figure 2. Developmental changes in behavior and reliability. Accuracy: The ability to correctly identify high arousal-negative (threat-related) emotions (fear and anger) improved with age in both the fMRI (***ai***) and full behavioral samples (***bi***). Only reliability of significant effects was assessed. RT: Time to identify an emotion decreased with age across all five emotion categories in both the fMRI (***aii***) and full behavioral samples (***bii***). ***c***, Reliability was measured as the probability of observing a significant correlation between accuracy and age or that between RT and age. ***ci***, Accuracy: Reliability of age-accuracy associations increased as sample size increased for conditions that showed significant age-related changes in accuracy in the fMRI (solid line) and the full behavioral (dashed line) samples. ***cii***, RT: Reliability of age-RT associations increased as sample size increased for all five emotion categories in the fMRI (solid line) and the full behavioral (dashed line) samples. In general, a sample size of 300–350 was needed to identify reliable (reliability > 0.7) age-accuracy association and that of 200 was needed to identify reliable age-RT association.

ing is not surprising given that the effect size during fear and anger conditions was comparable across the two samples. More-over, for sad and happy, the two emotion categories that showed a significant age effect only in the full behavioral sample, reliabil-ity increased with sample size but remained low (<0.5), even at a sample size of 700, reflecting the small effect size under these two

conditions. In addition, reliability of negative age-RT associations increased with sample size across all five emotion categories; and in general, a sample size of 150–200 was needed to identify reliable age-RT associations (Fig. 2*cii*). Results also demonstrate that age-related improvements in emotion identification were strongest for fear and anger, the two high arousal-negative emotions.

Behavioral and demographic differences between the neuroimaging and excluded samples

Finally, we examined whether participants in the fMRI ($N = 759$) and excluded ($N = 579$) samples are similar cross-sections of the overall population in terms of age, sex, and behavioral performance (accuracy and RT). The excluded sample included participants who did not meet our criteria for inclusion in fMRI data analysis but were included in the behavioral analysis. We found a significant difference in age between the fMRI and excluded samples ($t_{(1186)} = 11.68$, $p < 2.2e-16$), with participants in the fMRI sample (age: 16.20 ± 3.27 years) older than participants in the excluded sample (age: 13.99 ± 3.56 years). Moreover, sex and exclusion/inclusion were not independent ($\chi^2_{(1)} = 6.52$, $p = 0.011$), with higher female/male ratio in the fMRI sample (F/M = 419/340) than that in the excluded sample (F/M = 278/301). Last, we found significant differences in accuracy ($t_{(781)} = 8.83$, $p < 2.2e-16$) and RT ($t_{(998)} = -8.87$, $p < 2.2e-16$) between the two samples, with participants in the fMRI sample showing higher accuracy and faster response (accuracy: 0.92 ± 0.05; RT: 2.29 ± 0.36) than participants in the excluded sample (accuracy: 0.88 ± 0.10; RT: 2.51 ± 0.51). These findings are not surprising as children; and boys, in particular, usually tend to move more in the scanner (Yuan et al., 2009; Dosenbach et al., 2017; Engelhardt et al., 2017), so the excluded sample includes more boys and younger participants, leading to age and performance differences between the neuroimaging and excluded samples.

Meta-analysis of brain areas involved in emotion

To take a broad and inclusive view of cortical and subcortical brain regions involved in emotion perception, we proceeded in a two-step manner. First, we used the meta-analysis toolbox Neurosynth (Yarkoni et al., 2011) and the search term "emotion" to identify the most consistently reported cortical and subcortical regions implicated in emotion processing, including vmPFC, dmPFC, vlPFC, dlPFC, lOFC, IPL, SPL, sgACC, pgACC, dACC, PCC, pre-SMA, BLA, vAI, dAI, PI, hippocampus, and FFG (Fig. 1b; Table 3). Second, given recent theoretical focus on the salience network (Menon, 2011), specifically the subdivisions of the insula and the amygdala, we identified additional distinct functional subdivisions of the anterior insula (vAI) and amygdala (CMA) and additional subcortical regions, including the BNST and the NAc, thought to play distinct functional roles in emotion processing (LeDoux, 2007; Deen et al., 2011; Chang et al., 2013; Floresco, 2015; Lebow and Chen, 2016; Namkung et al., 2017). The brain areas identified by this analysis showed strong overlap with salience, central executive, and default mode networks (Seeley et al., 2007; Menon, 2011).

Modular structure of emotion-related circuits is established in childhood and remains developmentally stable

To investigate how the functional organization of emotion perception circuits change with age, we first identified the modular structure of interregional connectivity, computed using β series correlation (Rissman et al., 2004), for each emotion task condition. Next, we determined common patterns of modular structure across all participants for each emotion condition, which was used as a template for further analysis. Across all five emotion categories, we observed three consistent network modules corresponding to a FP module, a mPFC/PCC module, and an subcortical/posterior insula (SPI) module (Fig. 3a). This 3-module structure was stable across emotion categories for all values of gamma > 0.5; results are reported here for gamma $= 1$, the most commonly used value in most fMRI studies. The modularity

value Q for gamma $= 1$ was highly significant compared with empirical null networks with preserved weight and degree distributions ($p < 0.001$).

Next, we assessed main effects of age, emotion, and age $\times$ emotion interactions on the Jaccard overlap between individual and template modular structure. Results revealed a significant emotion effect ($F_{(4,3028)} = 3.656$, $p = 0.006$) but no significant effect of age or age $\times$ emotion interactions (p values >0.1), indicating that the overall modular structure for all emotion conditions was stable with age. Additional analysis using the Jaccard overlap between individual and adult-group modular structure revealed similar results. This was further confirmed by findings of modular structure within distinct age groups: children (8–12 years), adolescents (13–17 years), and adults (18–23 years) (Fig. 3b). Notably, this modular architecture maps closely onto the triple-network model (Menon, 2011). Finally, a similar 3-module structure was observed when using the 46 spatially extended clusters, demonstrating the robustness and generalizability of our finding with respect to Brainnetome functional brain parcellations (Fan et al., 2016).

Development of hubs underlying emotion perception

We next identified which nodes served as hubs and examined whether these change with age. Briefly, we identified hub regions for each individual, computed the probability of a brain region being a hub within a 1 year age window for each emotion category, and assessed the significance of hub probability by permutation. We found that hubs resided in the medial prefrontal nodes (pgACC, vmPFC, and dmPFC) of the mPFC/PCC module (Fig. 4a) and were stable over age across all five emotion categories (Fig. 4b,c).

Developmental changes in intramodule and intermodule connectivity

We examined main effects of age, emotion, and age $\times$ emotion interactions on intramodule and intermodule connectivity. Results revealed a significant main effect of age for intra-FP connectivity ($F_{(1,2692)} = 7.015$, $p = 0.008$) and for connectivity between FP and SPI ($F_{(1,2766)} = 5.253$, $p = 0.022$; Fig. 3c), indicating that intra-FP and FP-SPI connectivity increased with age. Results revealed a significant emotion effect in intra-FP ($F_{(4,3028)} = 7.487$, $p = 5.35e-06$), intra-SPI ($F_{(4,3028)} = 4.184$, $p = 0.002$), between FP and mPFC/PCC ($F_{(4,3028)} = 10.546$, $p = 1.75e-08$), between FP and SPI ($F_{(4,3028)} = 3.980$, $p = 0.003$), and between mPFC/PCC and SPI ($F_{(4,3028)} = 5.531$, $p = 1.9e-04$) connectivity (Fig. 3c,d). Age $\times$ emotion interactions are significant for intra-mPFC/PCC connectivity ($F_{(4,3028)} = 2.724$, $p = 0.028$). Further analysis revealed that intra-mPFC/PCC connectivity increased with age under fear ($F_{(1,754)} = 3.910$, $p = 0.048$) and anger ($F_{(1,754)} = 4.579$, $p = 0.033$) but not the other three emotion categories (p values > 0.1).

Developmental changes in interregional connectivity

Using the modular architecture identified in the previous sections, we examined main effects of age, emotion, and age $\times$ emotion interactions on interregional connectivity. All results are reported at $q < 0.05$, FDR corrected for multiple comparisons.
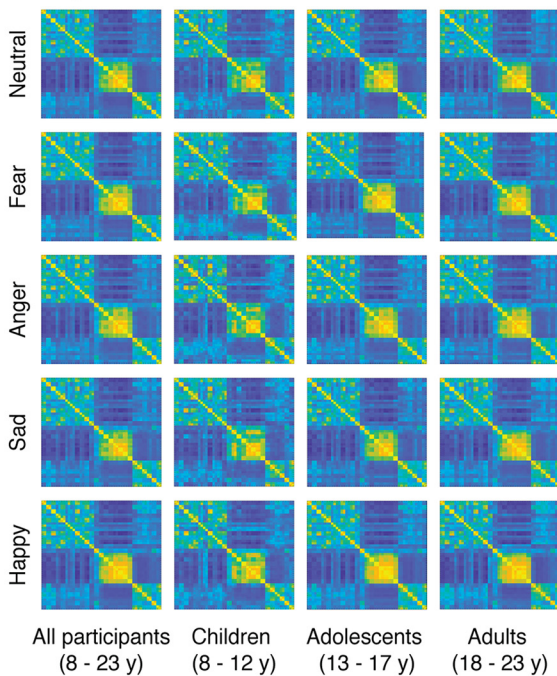
Main effect of age

We found age-related increases within FP and SPI modules, between FP and SPI modules, and between mPFC/PCC and SPI modules, including left dlPFC connectivity with right dAI, right vAI connectivity with right sgACC and bilateral PI, right BNST
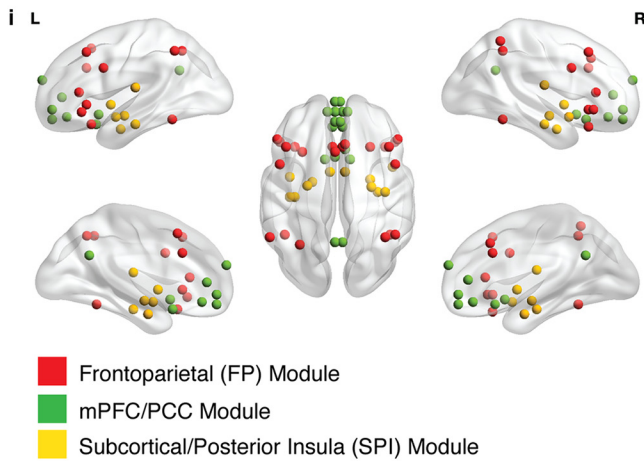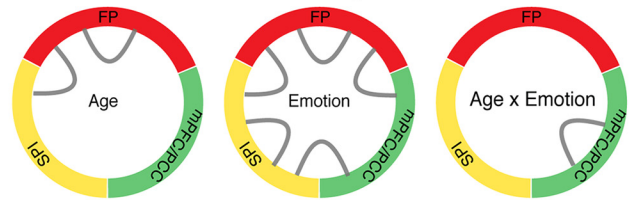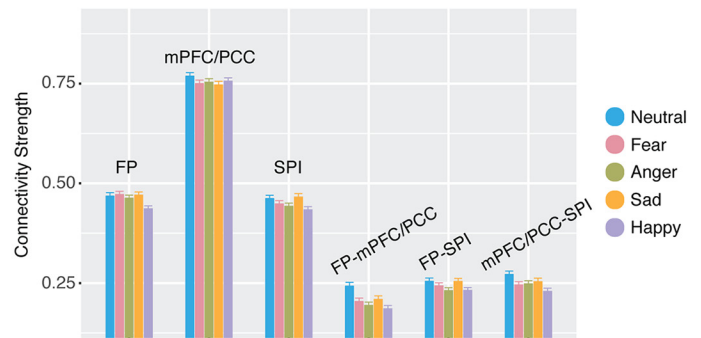
Figure 3. Main effects of age and emotion on modular organization and connectivity. ***ai***, Three functional modules were identified across all participants: the FP module (red), mPFC/PCC module (green), and SPI module (yellow). ***aii***, As shown in the group consistency matrix, the FP module (red border) includes bilateral dlPFC, vlPFC, dACC, pre-SMA, IPL, SPL, and FFG. The mPFC/PCC module (blue border) includes bilateral NAc, sgACC, pgACC, PCC, vmPFC, and dmPFC. The SPI module (yellow border) includes BLA, CMA, vAI, PI, BNST, and hippocampus (Hipp). Color bar represents the probability of two nodes being classified in the same module across participants. ***b***, As shown in the group consistency matrices, overall modular structure is stable across all five emotion categories and age groups (children 8–12 years, adolescents 13–17 years, adults 18–23 years). ***c***, Intra-FP connectivity and FP-SPI connectivity increased with age; intra-FP, intra-SPI, and all intermodule connectivity measures varied across emotions; age-related changes in intra-mPFC/PCC connectivity were modulated by emotion category. ***d***, Intramodule and intermodule connectivity varied across emotions. Data are mean ± SEM.

connectivity with right IPL and SPL, and right BLA connectivity with left PI (Fig. 5a). We found age-related decreases within mPFC/PCC and SPI modules, including right sgACC connectivity with left vmPFC and left CMA connectivity with right CMA and left PI. In addition, we found two links (left NAc connectivity with bilateral dACC), which showed significant quadratic age effects, consistent with findings in the literature regarding enhanced sensitivity of dopaminergic reward circuits in adolescence (Padmanabhan et al., 2011; Padmanabhan and Luna, 2014).

Main effect of emotion

We found widespread emotion effect, especially within FP module, between FP and mPFC/PCC modules, between FP and SPI modules, and between mPFC/PCC and SPI modules (Fig. 5a). Classification analysis revealed that it was sufficient to discriminate all pairs of emotion categories using all interregional links (Fig. 5c), with classification accuracy ranging from 56.2% to 67.2% and significantly higher than chance level (p values < 0.01) as assessed by permutation test.

a Nodal degree of circuit connectivity for each emotion category
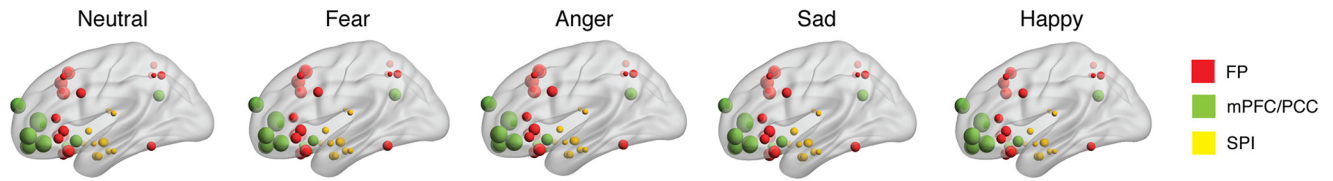
b Hub probability for each emotion category over development
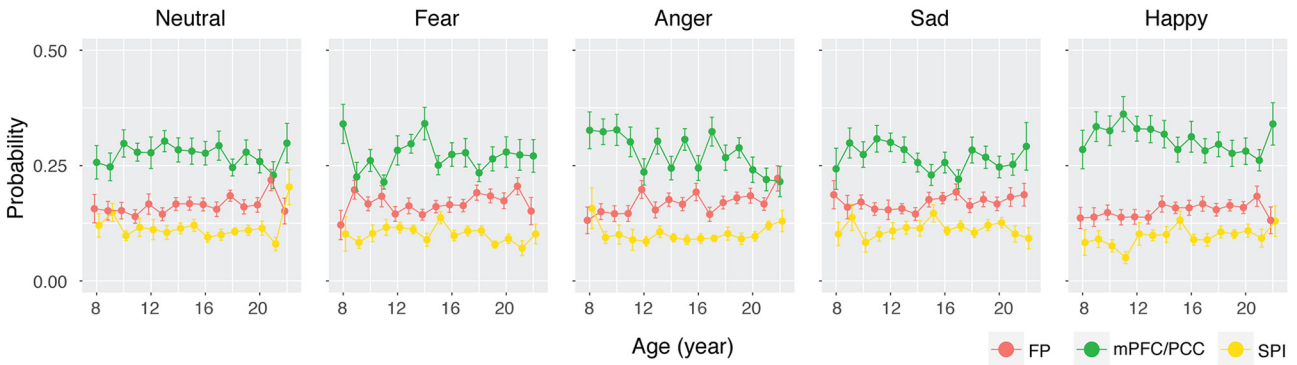
c Developmentally stable hubs in mPFC/PCC

Figure 4. Developmentally stable hubs in mPFC. **a**, ROI nodal degree for each emotion category. Node size is in proportion to integrated nodal degree measure. **b**, Probability measures averaged across all regions within each module at each age year. Overall, the mPFC/PCC module has a higher probability to be a functional hub across age than the FP and SPI modules. Data are mean ± SEM. **c**, Multiple regions of the mPFC/PCC module, including pgACC, vmPFC, and dmPFC of the mPFC/PCC module were identified as functional hub regions for all five emotion categories and were stable over development. Here, probability matrices are masked by the 95th percentile of empirical null distribution at each region and age year, with probability values ≤95th percentile set as 0. An age year includes all participants whose age was larger than or equal to the current age year and less than the next age year (e.g., age 8 includes all participants with age ≥8 and <9).

Age × emotion interactions
No age × emotion interactions were found.

Reliability of age-related changes
To examine the robustness of our age-related changes and the effect of sample size on reliability, we performed reliability anal-

ysis (Fig. 1d). Briefly, we randomly drew N (sample size) participants from the full sample and assessed age-related changes in the subsample using linear mixed-effect model. This was repeated 1000 times at each sample size across a wide range of sample sizes (50 ≤ N ≤ 400, increased by 50). Correlation be-

a Main effect of age and emotion on inter-regional connectivity (FDR corrected)
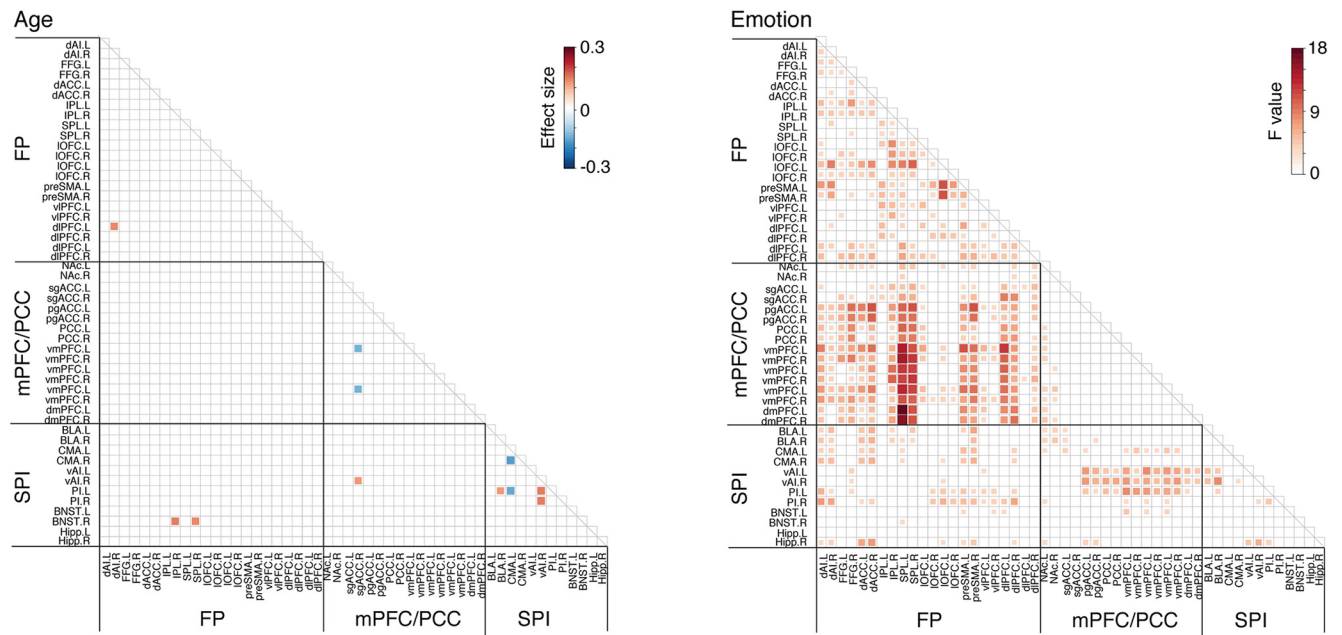
b Effect of sample size on reliability of age-related changes
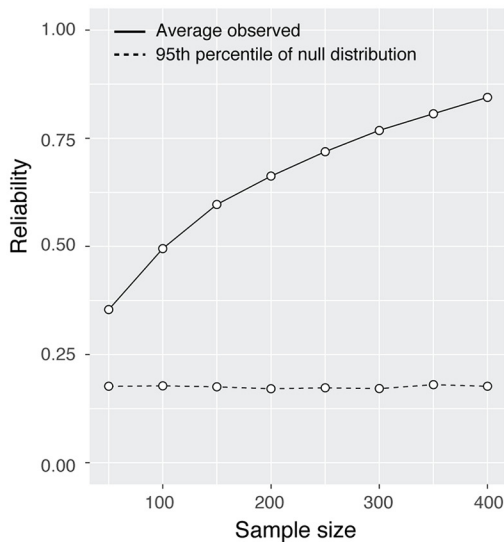
c Classification accuracy for differentiating connectivity patterns of emotions
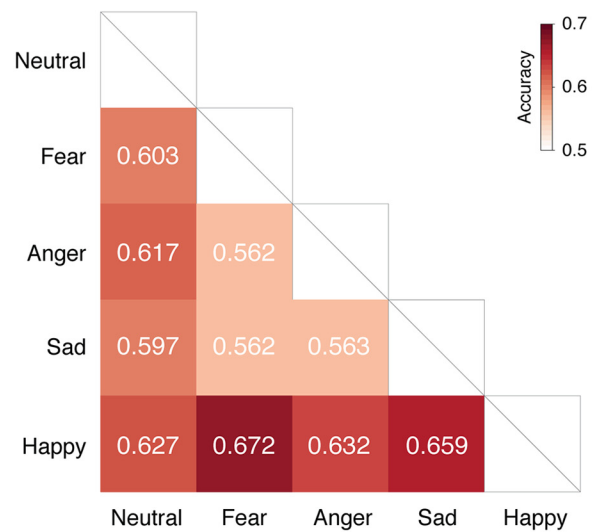
Figure 5. Main effects of age and emotion on regional connectivity. *a*, Connectivity within FP and SPI modules, between FP and SPI modules, and between mPFC/PCC and SPI modules increased with age, including left dlPFC connectivity with right dAI, right vAI connectivity with right sgACC and bilateral PI, right BNST connectivity with right IPL and SPL, and right BLA connectivity with left PI. Connectivity within mPFC/PCC and SPI modules decreased with age, including right sgACC connectivity with left vmPFC and left CMA connectivity with right CMA and left PI. Emotion effects are widespread, especially within FP module, between FP and mPFC/PCC modules, between FP and SPI modules, and between mPFC/PCC and SPI modules. *b*, Reliability of age-related changes in regional connectivity across all brain regions increased with sample size. A sample size of 50 is sufficient for reliability to surpass the 95th percentile of its corresponding empirical null distribution (dashed line), and a minimum sample size of 200 –250 was needed to identify reliable age-related changes. *c*, Classification analysis showed that it was sufficient to discriminate all pairs of emotion categories using all interregional links, with classification accuracy ranging from 56.2% to 67.2% and significantly higher than chance level as assessed by permutation test.

tween age effects from a subsample and that from the full sample was calculated and then averaged across subsamples at each sample size. The resulting average correlation was used to measure the reliability of overall developmental patterns across all interregional links. We found that reliability increased with sample size, that a minimum sample size of 50 was sufficient for reliability to surpass the 95th percentile of its corresponding empirical null distribution, and that a minimum sample size of 200 –250 was needed to identify reliable (reliability > 0.7) age-related

changes (Fig. 5*b*). These results highlight the robustness of our main age-related effects and indicate that large samples are needed to produce reliable findings.

Amygdala responses are invariant with age, whereas FP node responses decrease with age
We examined main effects of age, emotion, and age × emotion interactions on regional brain activity in each of the ROIs. All results reported here at *q* < 0.05, FDR corrected for multiple comparisons.

Main effect of age
We found significant age-related decreases in cognitive control regions, including bilateral pre-SMA, dACC and dlPFC, left SPL, right IPL, and right lOFC (Fig. 6ai).

Main effect of emotion
We found a significant emotion effect in all ROIs, except bilateral BNST (Fig. 6b).

Age × emotion interactions
No age × emotion interactions were found.

Reliability of age-related changes
We performed reliability analysis using the procedure same to that described in Developmental changes in interregional connectivity (Fig. 1d). We found that reliability increased with sample size, that a minimum sample size of 50–100 was sufficient for reliability to surpass the 95th percentile of its corresponding empirical null distribution, and that a minimum sample size of 150–200 was needed to identify reliable (reliability > 0.7) age-related changes (Fig. 6aii). These results highlight the robustness of our main age-related effects and indicate that large samples are needed to produce reliable findings.

Replicability of previous developmental findings related to amygdala circuitry
Given the inconsistencies with regards to developmental changes in amygdala recruitment and connectivity in the extant literature, we further examined age-related changes in the current large sample using precise anatomical coordinates reported in three previously published studies with similar emotion face recognition tasks, experimental contrasts, and age ranges (Table 4) (Gee et al., 2013; Kujawa et al., 2016; Wu et al., 2016).

Amygdala activity
Gee et al. (2013) reported age-related decreases in right amygdala activity during fearful face processing. We first included all five emotion categories in our linear mixed-effect model and did not identify age effect or age × emotion interactions in their reported amygdala region in the current PNC sample (F values < 2, p values > 0.1). We further examined age-related changes for each of the five emotion categories and still did not identify any age-related changes (Table 5). Kujawa et al. (2016) and Wu et al. (2016) reported no age-related changes in amygdala activity during identification of happy, fearful, and angry faces. When we included all five emotion categories in our linear mixed-effect model, we found no age-related changes or age × emotion interactions (F values < 1, p values > 0.1). When we included only fear, anger, and happy, the three emotion categories used in Kujawa et al. (2016) and Wu et al. (2016) studies, we found significant age-related increases ($F_{(1,1886)}$ = 4.357, p = 0.037, effect size = 0.076) and age × emotion interactions ($F_{(2,1514)}$ = 5.804, p = 0.003) in left amygdala using coordinates reported by Kujawa et al. (2016), and significant age-related increases ($F_{(1,1681)}$ = 4.962, p = 0.026, effect size = 0.081) in right amygdala using coordinates reported by Wu et al. (2016). We further examined age-related changes for each of the five emotion categories. Age-related increases were found in the left amygdala under fear, sad, and happy using coordinates reported by Kujawa et al. (2016) and in bilateral amygdala under fear, sad, and/or happy using coordinates reported by Wu et al. (2016) (Table 5). In summary, our replication results from the large sample are partially consistent with Kujawa et al. (2016) and Wu et al. (2016) findings and inconsistent with Gee et al. (2013) findings (see Table 7).
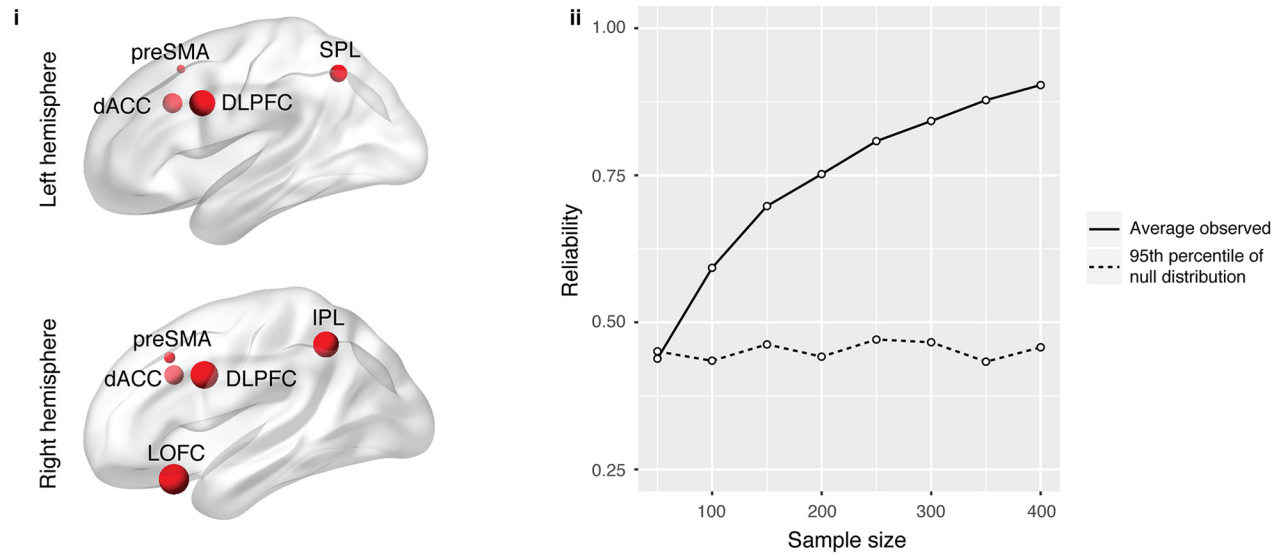
Amygdala connectivity
Gee et al. (2013) reported age-related decreases in the right amygdala-vmPFC connectivity during fearful face processing. We included five emotion categories in our linear mixed-effect model and did not identify age effect or age × emotion interactions in their reported amygdala-vmPFC regions in the current PNC sample (F values < 1, p values > 0.1). We further examined age-related changes for each of the five emotion categories and did not identify any age-related changes (Table 6). Kujawa et al. (2016) reported age-related decreases in bilateral amygdala-dACC connectivity during happy, angry, and fearful face identification in typically developing individuals. When we included five emotion categories in linear mixed-effect model, we found no age-related changes or age × emotion interactions using the reported coordinates (F values < 1, p values > 0.1). When we included only fear, anger, and happy, the three emotion categories used in the Kujawa et al. (2016) study, we found a significant age-related increases in right amygdala-dACC connectivity ($F_{(1,2011)}$ = 8.250, p = 0.004, effect size = 0.104). We further examined age-related changes for each of the five emotion categories. Age-related increases were found in left amygdala-dACC under fear and happy and in right amygdala-dACC under fear, anger, and sad (Table 6). Finally, Wu et al. (2016) reported age-related decreases in bilateral amygdala-mPFC/ACC during identification of fearful, angry, and happy faces. When we included all five emotion categories in linear mixed-effect model, we found no age effect or age × emotion interactions (F values < 2, p values > 0.1). When we included only fear, anger, and happy, the three emotion categories used in the Wu et al. (2016) study, we found significant age-related increases in right amygdala connectivity with left mPFC ($F_{(1,2041)}$ = 8.046, p = 0.005, effect size = 0.103) and right ACC ($F_{(1,2037)}$ = 6.749, p = 0.009, effect size = 0.094). We further examined age-related changes for each of the five emotion categories and found age-related increases in bilateral amygdala-mPFC/ACC connectivity under fear and/or anger (Table 6).

Amygdala connectivity (gPPI analysis)
Because previous developmental studies of amygdala connectivity have used PPI (Gee et al., 2013) or gPPI techniques (Kujawa et al., 2016; Wu et al., 2016), we conducted additional analyses using a similar approach. When we included the five emotion categories in linear mixed-effect model, we did not find a main effect of age, or age × emotion interactions in analyses using coordinates from the three studies (F values < 1.5, p values > 0.1). When we included only fear, anger and happy in linear mixed-effect model, we did not identify age effect or age × emotion interactions in any of the amygdala connectivity using coordinates from Kujawa et al. (2016) and Wu et al. (2016) studies (F values < 2.2, p values > 0.1). Finally, we examined age-related changes for each emotion category. Results revealed no age effect using coordinates reported in Gee et al. (2013) study (F values < 2, p values > 0.1) but significant age-related increases in bilateral amygdala-dACC connectivity under fear (left: $F_{(1,754)}$ = 5.602, p = 0.018, effect size = 0.088; right: $F_{(1,754)}$ = 4.605, p = 0.032, effect size = 0.080) using coordinates reported by Kujawa et al. (2016) as well as in left amygdala-mPFC connectivity under fear ($F_{(1,754)}$ = 7.041, p = 0.008, effect size = 0.099) using coordinates reported by Wu et al. (2016). These results are consistent with those derived using β time series analysis, as described in Amygdala connectivity.

Together, these results point to poor replicability of previous developmental findings on amygdala connectivity associated with emotion processing (Table 7).

Figure 6. Main effects of age and emotion on regional activity. ***ai***, Brain regions showing age-related changes. Activity in cognitive control regions, including bilateral pre-SMA, dACC, dlPFC, left SPM, right IPL, and right lOFC, decreased with age. Node size was scaled to present effect size of age-related decrease. ***aii***, Reliability of age-related changes across all brain regions. A minimum sample size of 50 –100 was needed for reliability to surpass the 95th percentile of its corresponding empirical null distribution (dashed line), and a minimum sample size of 50 –100 was needed to identify reliable (correlation > 0.7) age-related changes. ***bi***, Brain regions showing emotion effects. Node size was scaled to present *F* value of emotion effect. ***bii***, Activity in all brain regions, except bilateral BNST, varied across emotion category.

Table 5. Age-related changes in amygdala activity in the PNC sample using coordinates from previous studies[a]

Region	Emotion	$F_{(1,754)}$	p	Effect size
Right amygdala in Gee et al. (2013) ↓	Neutral	1.121	0.290	−0.040
	Fear	0.228	0.633	0.018
	Anger	0.174	0.677	−0.016
	Sad	1.965	0.161	0.052
	Happy	0.109	0.741	−0.012
Left amygdala in Kujawa et al. (2016)[b]	Neutral	0.953	0.329	0.036
	Fear	12.617	0.0004	0.132
	Anger	1.844	0.175	0.051
	Sad	15.140	0.0001	0.144
	Happy	9.020	0.003	0.111
Right amygdala in Kujawa et al. (2016)[b]	Neutral	0.566	0.452	0.028
	Fear	2.650	0.104	0.061
	Anger	0.123	0.726	0.013
	Sad	2.801	0.095	0.063
	Happy	0.726	0.394	0.032
Left amygdala (AAL) in Wu et al. (2016)[b]	Neutral	1.583	0.209	0.047
	Fear	15.796	7.73e-05	0.147
	Anger	4.571	0.033	0.080
	Sad	20.755	6.08e-06	0.168
	Happy	10.009	0.002	0.117
Right amygdala (AAL) in Wu et al. (2016)[b]	Neutral	0.277	0.599	0.020
	Fear	5.741	0.017	0.089
	Anger	1.946	0.163	0.052
	Sad	5.784	0.016	0.089
	Happy	1.553	0.213	0.046

[a]Effect size is measured by the standardized partial coefficient of the age term. AAL, Automated anatomical labeling; ↓, developmental decreases in amygdala with mPFC was reported in the cited study. AAL indicates that the amygdala ROI was derived based on AAL atlas.

[b]No significant developmental changes were found in the cited study.

Discussion

A large sample of children, adolescents, and young adults with high quality emotion perception task fMRI data from the PNC (Satterthwaite et al., 2014) afforded us an unprecedented opportunity to address critical gaps in our knowledge regarding the development of human emotion identification circuits. We used a multipronged analytical strategy to investigate the development of functional circuits associated with perception and identification of five distinct emotion categories: fear, anger, sadness, happiness, and neutral. Our analysis disentangled aspects of emotion-related brain circuitry that were stable over development and those that changed with age. We found a developmentally stable modular architecture anchored in the salience, default mode, and central executive networks, with hubs in the mPFC and emotion-related reconfiguration of the salience network. Developmental changes were most prominent in FP circuits important for salience detection and cognitive control. Crucially, reliability analyses provided quantitative evidence for the robustness of our findings. Our findings provide a new template for investigation of emotion processing in the developing brain.

Maturation of emotion identification and specificity with respect to high arousal-negative emotions

Behaviorally, facial emotion identification emerges early in infancy with continued improvements through childhood and adolescence (Batty and Taylor, 2006; Thomas et al., 2007; Rodger et al., 2015; Leitzke and Pollak, 2016; Theurel et al., 2016). We found that, although emotion identification accuracy was high for all five emotion categories, even in the youngest individuals (age 8), identification of fear and anger emotion categories improved significantly with age, consistent with previous reports (Gee et al., 2013; Kujawa et al., 2016; Wu et al., 2016). These

Table 6. Age-related changes in amygdala connectivity in the PNC sample using coordinates from previous studies[a]

Connectivity	Emotion	$F_{(1,754)}$	p	Effect size
AMY.R-vmPFC.R in Gee et al. (2013) ↓	Neutral	0.210	0.647	0.017
	Fear	3.497	0.062	0.069
	Anger	1.880	0.171	0.051
	Sad	0.963	0.327	0.036
	Happy	0.288	0.591	0.020
AMY.L-dACC.L in Kujawa et al. (2016) ↓	Neutral	0.008	0.929	−0.003
	Fear	8.921	0.003	0.109
	Anger	2.569	0.109	0.059
	Sad	0.388	0.533	0.023
	Happy	4.596	0.032	0.079
AMY.R-dACC.R in Kujawa et al. (2016) ↓	Neutral	0.214	0.644	0.017
	Fear	6.602	0.010	0.094
	Anger	7.661	0.006	0.101
	Sad	4.420	0.036	0.077
	Happy	3.232	0.073	0.066
AMY.L-mPFC.L in Wu et al. (2016) ↓	Neutral	0.012	0.914	−0.004
	Fear	4.139	0.042	0.075
	Anger	1.198	0.274	0.040
	Sad	0.003	0.958	0.002
	Happy	1.199	0.274	0.040
AMY.R-mPFC.L in Wu et al. (2016) ↓	Neutral	0.334	0.563	0.021
	Fear	3.712	0.054	0.071
	Anger	7.566	0.006	0.101
	Sad	1.333	0.249	0.004
	Happy	2.743	0.098	0.061
AMY.L-ACC.L in Wu et al. (2016) ↓	Neutral	0.019	0.890	0.005
	Fear	4.351	0.037	0.076
	Anger	3.000	0.084	0.064
	Sad	0.602	0.438	0.028
	Happy	2.769	0.097	0.061
AMY.R-ACC.R in Wu et al. (2016) ↓	Neutral	0.013	0.909	0.004
	Fear	4.551	0.033	0.078
	Anger	5.901	0.015	0.089
	Sad	4.496	0.034	0.078
	Happy	1.261	0.262	0.041

[a]Effect size is measured by the standardized partial coefficient of age term. ↓, Developmental decrease in amygdala-mPFC connectivity (reported in the cited study).

findings suggest that high arousal-negative ("threat-related") emotions show more prominent differentiation, allowing for further sharpening of emotion concepts with age (Widen, 2013; Nook et al., 2017).

Architecture and developmental stability of emotion circuitry

Network analysis of interregional connectivity provided novel evidence for a developmentally stable architecture of emotion-related circuitry. We found three distinct communities (FP, mPFC/PCC, and SPI) that shared similarities with canonical intrinsic salience, central executive, and default mode networks (Menon, 2011), with some key differences.

We found significant reconfiguration of the salience network during emotion identification, including segregation of individual insula subdivisions, amygdala, and dACC, into separate modules. The dorsal AI and the dACC grouped with the FP community of cognitive control systems, whereas the ventral AI, PI, and amygdala grouped with the SPI module of subcortical/limbic systems. This is noteworthy because the dorsal, ventral, anterior, and posterior subdivisions of the insula play distinct and integrative roles in cognitive-affective aspects of emotion perception (Craig, 2009; Menon and Uddin, 2010; Uddin et al., 2014). Findings provide novel evidence that the overall functional archi-

Table 7. Replicability of previous developmental findings in amygdala activity and connectivity underlying face emotion perception

Study	Findings	Replicability	
		Amygdala activity	Amygdala connectivity
Gee et al. (2013)	Amygdala activity decreased with age; amygdala-vmPFC connectivity decreased with age	No	No
Kujawa et al. (2016)	No age-related effect in amygdala activity; amygdala-dACC connectivity decreased with age in TD but increased with age in AD for all emotions	Partially yes	No
Wu et al. (2016)	No age-related effect in amygdala activity; amygdala-ACC/mPFC connectivity decreased with age for all emotions	Partially yes	No

tecture of emotion identification is largely stable across development and different categories of emotion.

Multiple mPFC regions form a developmentally stable core of emotion circuitry

Network analysis also revealed developmentally stable hubs in mPFC regions pgACC, vmPFC, and dmPFC across all five emotion categories. These regions play central roles in socioemotional processing (Phillips et al., 2003; Ochsner et al., 2004; Peelen et al., 2010) The vmPFC and dmPFC are thought to be important for self-referential processing and mentalizing (Ochsner et al., 2012; Dixon et al., 2017), whereas the pgACC is important for awareness of one's own emotions (Lane et al., 2015; Dixon et al., 2017). Last, our results are in line with the theoretical framework that the mPFC plays a central role in emotion categorization by tuning other brain systems to distinct emotion categories (Heberlein et al., 2008; Roy et al., 2012; Barrett and Satpute, 2013; Satpute et al., 2013, 2016).

Surprisingly, while the PCC was part of the same module as the mPFC, it was not a hub region, despite being consistently identified as hub in intrinsic functional connectivity studies (Buckner et al., 2008; Fransson and Marrelec, 2008; Greicius et al., 2009; Andrews-Hanna et al., 2010). These key differences between intrinsic and task-based network organization argue for the critical importance of characterizing task-dependent brain circuits.

Modular and regional connectivity changes are largely driven by emotion category and are stable with age

Anchored on a stable modular architecture, we found significant differences in intramodule and intermodule and regional connectivity associated with emotion category, and to a lesser degree age. Prominent differences between emotion categories provide evidence that distinct emotions, although anchored in a common modular architecture, may be represented differently in the brain at the network level (Kragel and LaBar, 2015; Saarimäki et al., 2016). Our support vector machine-based classification provides evidence of unique brain connectivity patterns elicited by each emotion category and indicate differential patterns of distributed network connectivity patterns involving the same regions (Barrett, 2017).

While module connectivity was largely developmentally stable, two findings of age-related changes are noteworthy. First, significant age-related increases in FP intramodule connectivity point to continued maturation of FP control systems underlying emotion categorization over development (Ahmed et al., 2015; Guyer et al., 2016). Second, age-related increases in mPFC/PCC intramodule connectivity during identification of fearful and angry faces suggest that maturation of midline DMN systems underlie continued developmental improvements in perception of threat and high emotion-arousal emotions. Last, significant developmental changes in regional connectivity centered predominantly on the insula, a core node of the salience network, highlighting a central and previously underrecognized role of the insula in the development of emotion perception and identification.

Developmental changes in amygdala reactivity and amygdala-PFC circuits

Surprisingly, we did not detect developmental changes in the amygdala and its functional circuitry, given that amygdala and its connectivity with the mPFC have been the focus of developmental studies of emotion processing, but with contradictory findings (for a summary, see Table 1). To address these discrepancies, we examined amygdala activity and amygdala-mPFC connectivity in our PNC cohort using specific amygdala seed and mPFC target regions from three published studies with similar emotion face recognition tasks, experimental contrasts, and age ranges (Gee et al., 2013; Kujawa et al., 2016; Wu et al., 2016). We were unable to replicate the findings of these previous studies, suggesting that small samples, and not differences in amygdala and mPFC coordinates, are a major contributor to the discrepancies between the current and previous studies. These findings suggest that a narrow focus on amygdala-mPFC connectivity has precluded characterization of other important brain regions, most notably the insula and other PFC regions, underlying the development of emotion perception and categorization.

Further studies are needed to determine the extent to which differences in experimental design contributed to inconsistent patterns of developmental changes in amygdala activity and connectivity reported in previous studies. Critically, addressing this question will require much larger developmental samples than used in previous studies, along with stability and reliability analyses as in the present study.

Neutral faces engage emotion circuitry in the context of emotion identification

Another surprising finding of our study was that neutral faces showed distributed and strong levels of activity and connectivity across the three modules. Disambiguation of neural faces from other emotions would require perception, evaluation, and decision-making and, as our data indicate, engage emotion perception and categorization circuits. Moreover, the classification rates for neutral faces were no different from those between the other emotions. These findings suggest that previous findings with neutral faces as a control condition in emotion identification tasks should be reassessed and interpreted with caution. This is also an important consideration for future studies of emotion identification.

Reliability analyses highlight robustness of findings

Lack of replication arising from small sample sizes is a significant problem in neuroscience research (Button et al., 2013; Szucs and Ioannidis, 2017). We examined the robustness of our findings by assessing reliability of findings at different sample sizes (Schaer et al., 2015; Moser et al., 2018; Turner et al., 2018). We found that sample sizes of ~200–300 allowed us to achieve a reliability of ≥0.70, providing strong evidence for the role of sample size in generating potentially reproducible findings and that the modest sample sizes of 25–65 used in previous studies have low reliability (<0.5).

It is also important to note that logistical challenges often preclude individual research groups from collecting large sam-

ples of data, which can significantly impede scientific progress (Nee, 2019). The recent curation and sharing of large-scale neurodevelopmental datasets, such as the PNC, provide a promising solution to this challenge and afford the opportunity to begin reconciliation of mixed findings in the extant literature.

In conclusion, emotion identification and categorization are fundamental to affective and cognitive development and involve integration of multiple widely distributed brain systems. We used quantitatively rigorous procedures that span multiple levels of regional, circuit, and network analysis and provide robust evidence for shared, unique, stable, and changing neural fingerprints of emotion perception development from childhood to adulthood. We stress the importance of large sample sizes and a systems-wide brain networks approach for understanding the development of human emotion perception.

References

Ahmed SP, Bittencourt-Hewitt A, Sebastian CL (2015) Neurocognitive bases of emotion regulation development in adolescence. Dev Cogn Neurosci 15:11–25.

Andrews-Hanna JR, Reidler JS, Sepulcre J, Poulin R, Buckner RL (2010) Functional-anatomic fractionation of the brain's default network. Neuron 65:550–562.

Avery SN, Clauss JA, Winder DG, Woodward N, Heckers S, Blackford JU (2014) BNST neurocircuitry in humans. Neuroimage 91:311–323.

Barrett LF (2017) The theory of constructed emotion: an active inference account of interoception and categorization. Soc Cogn Affect Neurosci 12:1–23.

Barrett LF, Satpute AB (2013) Large-scale brain networks in affective and social neuroscience: towards an integrative functional architecture of the brain. Curr Opin Neurobiol 23:361–372.

Bates D, Machler M, Bolker BM, Walker SC (2015) Fitting linear mixed-effects models using lme4. J Stat Softw 67:1–48.

Batty M, Taylor MJ (2006) The development of emotional face processing during childhood. Dev Sci 9:207–220.

Blackford JU, Pine DS (2012) Neural substrates of childhood anxiety disorders: a review of neuroimaging findings. Child Adolesc Psychiatr Clin North Am 21:501–525.

Blakemore SJ (2008) The social brain in adolescence. Nat Rev Neurosci 9:267–277.

Blondel VD, Guillaume JL, Lambiotte R, Lefebvre E (2008) Fast unfolding of communities in large networks. J Statist Mech: Theory Exp 10:P10008.

Braunstein LM, Gross JJ, Ochsner KN (2017) Explicit and implicit emotion regulation: a multi-level framework. Soc Cogn Affect Neurosci 12:1545–1557.

Buckner RL, Andrews-Hanna JR, Schacter DL (2008) The brain's default network: anatomy, function, and relevance to disease. Ann N Y Acad Sci 1124:1–38.

Buhle JT, Silvers JA, Wager TD, Lopez R, Onyemekwu C, Kober H, Weber J, Ochsner KN (2014) Cognitive reappraisal of emotion: a meta-analysis of human neuroimaging studies. Cereb Cortex 24:2981–2990.

Button KS, Ioannidis JP, Mokrysz C, Nosek BA, Flint J, Robinson ES, Munafò MR (2013) Power failure: why small sample size undermines the reliability of neuroscience. Nat Rev Neurosci 14:365–376.

Chang LJ, Yarkoni T, Khaw MW, Sanfey AG (2013) Decoding the role of the insula in human cognition: functional parcellation and large-scale reverse inference. Cereb Cortex 23:739–749.

Craig AD (2009) How do you feel—now? The anterior insula and human awareness. Nat Rev Neurosci 10:59–70.

Deen B, Pitskel NB, Pelphrey KA (2011) Three systems of insular functional connectivity identified with cluster analysis. Cereb Cortex 21:1498–1506.

Dixon ML, Thiruchselvam R, Todd R, Christoff K (2017) Emotion and the prefrontal cortex: an integrative review. Psychol Bull 143:1033–1081.

Dosenbach NU, Koller JM, Earl EA, Miranda-Dominguez O, Klein RL, Van AN, Snyder AZ, Nagel BJ, Nigg JT, Nguyen AL, Wesevich V, Greene DJ, Fair DA (2017) Real-time motion analytics during brain MRI improve data quality and reduce costs. Neuroimage 161:80–93.

Dosenbach NU, Fair DA, Miezin FM, Cohen AL, Wenger KK, Dosenbach RA, Fox MD, Snyder AZ, Vincent JL, Raichle ME, Schlaggar BL, Petersen SE

(2007) Distinct brain networks for adaptive and stable task control in humans. Proc Natl Acad Sci U S A 104:11073–11078.

Dwyer DB, Harrison BJ, Yücel M, Whittle S, Zalesky A, Pantelis C, Allen NB, Fornito A (2014) Large-scale brain network dynamics supporting adolescent cognitive control. J Neurosci 34:14096–14107.

Eickhoff SB, Stephan KE, Mohlberg H, Grefkes C, Fink GR, Amunts K, Zilles K (2005) A new SPM toolbox for combining probabilistic cytoarchitectonic maps and functional imaging data. Neuroimage 25:1325–1335.

Eklund A, Nichols TE, Knutsson H (2016) Cluster failure: why fMRI inferences for spatial extent have inflated false-positive rates. Proc Natl Acad Sci U S A 113:7900–7905.

Engelhardt LE, Roe MA, Juranek J, DeMaster D, Harden KP, Tucker-Drob EM, Church JA (2017) Children's head motion during fMRI tasks is heritable and stable over time. Dev Cogn Neurosci 25:58–68.

Etkin A (2010) Functional neuroanatomy of anxiety: a neural circuit perspective. Curr Top Behav Neurosci 2:251–277.

Fan L, Li H, Zhuo J, Zhang Y, Wang J, Chen L, Yang Z, Chu C, Xie S, Laird AR, Fox PT, Eickhoff SB, Yu C, Jiang T (2016) The human Brainnetome atlas: a new brain atlas based on connectional architecture. Cereb Cortex 26:3508–3526.

Floresco SB (2015) The nucleus accumbens: an interface between cognition, emotion, and action. Annu Rev Psychol 66:25–52.

Fortunato S, Barthélemy M (2007) Resolution limit in community detection. Proc Natl Acad Sci U S A 104:36–41.

Fransson P, Marrelec G (2008) The precuneus/posterior cingulate cortex plays a pivotal role in the default mode network: evidence from a partial correlation network analysis. Neuroimage 42:1178–1184.

Fusar-Poli P, Placentino A, Carletti F, Landi P, Allen P, Surguladze S, Benedetti F, Abbamonte M, Gasparotti R, Barale F, Perez J, McGuire P, Politi P (2009) Functional atlas of emotional faces processing: a voxel-based meta-analysis of 105 functional magnetic resonance imaging studies. J Psychiatry Neurosci 34:418–432.

Gee DG, Humphreys KL, Flannery J, Goff B, Telzer EH, Shapiro M, Hare TA, Bookheimer SY, Tottenham N (2013) A developmental shift from positive to negative connectivity in human amygdala-prefrontal circuitry. J Neurosci 33:4584–4593.

Greicius MD, Krasnow B, Reiss AL, Menon V (2003) Functional connectivity in the resting brain: a network analysis of the default mode hypothesis. Proc Natl Acad Sci U S A 100:253–258.

Greicius MD, Supekar K, Menon V, Dougherty RF (2009) Resting-state functional connectivity reflects structural connectivity in the default mode network. Cereb Cortex 19:72–78.

Guyer AE, Silk JS, Nelson EE (2016) The neurobiology of the emotional adolescent: from the inside out. Neurosci Biobehav Rev 70:74–85.

Gyurak A, Gross JJ, Etkin A (2011) Explicit and implicit emotion regulation: a dual-process framework. Cogn Emot 25:400–412.

Heberlein AS, Padon AA, Gillihan SJ, Farah MJ, Fellows LK (2008) Ventromedial frontal lobe plays a critical role in facial emotion recognition. J Cogn Neurosci 20:721–733.

Ioannidis JP (2005) Why most published research findings are false. PLoS Med 2:e124.

Kober H, Barrett LF, Joseph J, Bliss-Moreau E, Lindquist K, Wager TD (2008) Functional grouping and cortical-subcortical interactions in emotion: a meta-analysis of neuroimaging studies. Neuroimage 42:998–1031.

Kohn N, Eickhoff SB, Scheller M, Laird AR, Fox PT, Habel U (2014) Neural network of cognitive emotion regulation: an ALE meta-analysis and MACM analysis. Neuroimage 87:345–355.

Kragel PA, LaBar KS (2015) Multivariate neural biomarkers of emotional states are categorically distinct. Soc Cogn Affect Neurosci 10:1437–1448.

Kujawa A, Wu M, Klumpp H, Pine DS, Swain JE, Fitzgerald KD, Monk CS, Phan KL (2016) Altered development of amygdala-anterior cingulate cortex connectivity in anxious youth and young adults. Biol Psychiatry Cogn Neurosci Neuroimaging 1:345–352.

Lancichinetti A, Fortunato S (2012) Consensus clustering in complex networks. Sci Rep 2:336.

Lane RD, Weihs KL, Herring A, Hishaw A, Smith R (2015) Affective agnosia: expansion of the alexithymia construct and a new opportunity to integrate and extend Freud's legacy. Neurosci Biobehav Rev 55:594–611.

Lebow MA, Chen A (2016) Overshadowed by the amygdala: the bed nucleus of the stria terminalis emerges as key to psychiatric disorders. Mol Psychiatry 21:450–463.

LeDoux J (2007) The amygdala. Curr Biol 17:R868–R874.

Leitzke BT, Pollak SD (2016) Developmental changes in the primacy of facial cues for emotion recognition. Dev Psychol 52:572–581.

Lv Q, Yang L, Li G, Wang Z, Shen Z, Yu W, Jiang Q, Hou B, Pu J, Hu H, Wang Z (2016) Large-scale persistent network reconfiguration induced by ketamine in anesthetized monkeys: relevance to mood disorders. Biol Psychiatry 79:765–775.

Menon V (2011) Large-scale brain networks and psychopathology: a unifying triple network model. Trends Cogn Sci 15:483–506.

Menon V (2013) Developmental pathways to functional brain networks: emerging principles. Trends Cogn Sci 17:627–640.

Menon V, Uddin LQ (2010) Saliency, switching, attention and control: a network model of insula function. Brain Struct Funct 214:655–667.

Moser DA, Doucet GE, Lee WH, Rasgon A, Krinsky H, Leibu E, Ing A, Schumann G, Rasgon N, Frangou S (2018) Multivariate associations among behavioral, clinical, and multimodal imaging phenotypes in patients with psychosis. JAMA Psychiatry 75:386–395.

Namkung H, Kim SH, Sawa A (2017) The insula: an underestimated brain area in clinical neuroscience, psychiatry, and neurology. Trends Neurosci 40:200–207.

Nee DE (2019) fMRI replicability depends upon sufficient individual-level data. Commun Biol 2:130.

Newman ME (2004) Fast algorithm for detecting community structure in networks. Phys Rev E Stat Nonlin Soft Matter Phys 69:066133.

Newman ME (2006) Modularity and community structure in networks. Proc Natl Acad Sci U S A 103:8577–8582.

Nook EC, Sasse SF, Lambert HK, McLaughlin KA, Somerville LH (2017) Increasing verbal knowledge mediates development of multidimensional emotion representations. Nat Hum Behav 1:881–889.

Ochsner KN, Knierim K, Ludlow DH, Hanelin J, Ramachandran T, Glover G, Mackey SC (2004) Reflecting upon feelings: an fMRI study of neural systems supporting the attribution of emotion to self and other. J Cogn Neurosci 16:1746–1772.

Ochsner KN, Silvers JA, Buhle JT (2012) Functional imaging studies of emotion regulation: a synthetic review and evolving model of the cognitive control of emotion. Ann N Y Acad Sci 1251:E1–E24.

Padmanabhan A, Luna B (2014) Developmental imaging genetics: linking dopamine function to adolescent behavior. Brain Cogn 89:27–38.

Padmanabhan A, Geier CF, Ordaz SJ, Teslovich T, Luna B (2011) Developmental changes in brain function underlying the influence of reward processing on inhibitory control. Dev Cogn Neurosci 1:517–529.

Paulus MP, Stein MB (2006) An insular view of anxiety. Biol Psychiatry 60:383–387.

Peelen MV, Atkinson AP, Vuilleumier P (2010) Supramodal representations of perceived emotions in the human brain. J Neurosci 30:10127–10134.

Phillips ML, Drevets WC, Rauch SL, Lane R (2003) Neurobiology of emotion perception: II. Implications for major psychiatric disorders. Biol Psychiatry 54:515–528.

Power JD, Fair DA, Schlaggar BL, Petersen SE (2010) The development of human functional brain networks. Neuron 67:735–748.

Qin S, Young CB, Supekar K, Uddin LQ, Menon V (2012) Immature integration and segregation of emotion-related brain circuitry in young children. Proc Natl Acad Sci U S A 109:7941–7946.

Qin S, Young CB, Duan X, Chen T, Supekar K, Menon V (2014) Amygdala subregional structure and intrinsic functional connectivity predicts individual differences in anxiety during early childhood. Biol Psychiatry 75:892–900.

Rissman J, Gazzaley A, D'Esposito M (2004) Measuring functional connectivity during distinct stages of a cognitive task. Neuroimage 23:752–763.

Rodger H, Vizioli L, Ouyang X, Caldara R (2015) Mapping the development of facial expression recognition. Dev Sci 18:926–939.

Roy M, Shohamy D, Wager TD (2012) Ventromedial prefrontal-subcortical systems and the generation of affective meaning. Trends Cogn Sci 16:147–156.

Rubin TN, Koyejo O, Gorgolewski KJ, Jones MN, Poldrack RA, Yarkoni T (2017) Decoding brain activity using a large-scale probabilistic functional-anatomical atlas of human cognition. Plos Comput Biol 13:e1005649.

Rubinov M, Sporns O (2010) Complex network measures of brain connectivity: uses and interpretations. Neuroimage 52:1059–1069.

Saarimäki H, Gotsopoulos A, Jääskelainen IP, Lampinen J, Vuilleumier P, Hari R, Sams M, Nummenmaa L (2016) Discrete neural signatures of basic emotions. Cereb Cortex 26:2563–2573.

Satpute AB, Shu J, Weber J, Roy M, Ochsner KN (2013) The functional neural architecture of self-reports of affective experience. Biol Psychiatry 73:631–638.

Satpute AB, Nook EC, Narayanan S, Shu J, Weber J, Ochsner KN (2016) Emotions in "Black and White" or shades of gray? How we think about emotion shapes our perception and neural representation of emotion. Psychol Sci 27:1428–1442.

Satterthwaite TD, Elliott MA, Ruparel K, Loughead J, Prabhakaran K, Calkins ME, Hopson R, Jackson C, Keefe J, Riley M, Mentch FD, Sleiman P, Verma R, Davatzikos C, Hakonarson H, Gur RC, Gur RE (2014) Neuroimaging of the Philadelphia neurodevelopmental cohort. Neuroimage 86:544–553.

Schaer M, Kochalka J, Padmanabhan A, Supekar K, Menon V (2015) Sex differences in cortical volume and gyrification in autism. Mol Autism 6:42.

Seeley WW, Menon V, Schatzberg AF, Keller J, Glover GH, Kenna H, Reiss AL, Greicius MD (2007) Dissociable intrinsic connectivity networks for salience processing and executive control. J Neurosci 27:2349–2356.

Sperduti M, Makowski D, Arcangeli M, Wantzen P, Zalla T, Lemaire S, Dokic J, Pelletier J, Piolino P (2017) The distinctive role of executive functions in implicit emotion regulation. Acta Psychol (Amst) 173:13–20.

Sporns O, Betzel RF (2016) Modular brain networks. Annu Rev Psychol 67:613–640.

Szucs D, Ioannidis JP (2017) Empirical assessment of published effect sizes and power in the recent cognitive neuroscience and psychology literature. PLoS Biol 15:e2000797.

Theurel A, Witt A, Malsert J, Lejeune F, Fiorentini C, Barisnikov K, Gentaz E (2016) The integration of visual context information in facial emotion recognition in 5- to 15-year-olds. J Exp Child Psychol 150:252–271.

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.

Tupak SV, Dresler T, Guhn A, Ehlis AC, Fallgatter AJ, Pauli P, Herrmann MJ (2014) Implicit emotion regulation in the presence of threat: neural and autonomic correlates. Neuroimage 85:372–379.

Turner BO, Paul EJ, Miller MB, Barbey AK (2018) Small sample sizes reduce the replicability of task-based fMRI studies. Commun Biol 1:62.

Uddin LQ, Kinnison J, Pessoa L, Anderson ML (2014) Beyond the tripartite cognition-emotion-interoception model of the human insular cortex. J Cogn Neurosci 26:16–27.

Vink M, Derks JM, Hoogendam JM, Hillegers M, Kahn RS (2014) Functional differences in emotion processing during adolescence and early adulthood. Neuroimage 91:70–76.

Widen SC (2013) Children's interpretation of facial expressions: the long path from valence-based to specific discrete categories. Emotion Rev 5:72–77.

Wolf RC, Herringa RJ (2016) Prefrontal-amygdala dysregulation to threat in pediatric posttraumatic stress disorder. Neuropsychopharmacology 41:822–831.

Wu M, Kujawa A, Lu LH, Fitzgerald DA, Klumpp H, Fitzgerald KD, Monk CS, Phan KL (2016) Age-related changes in amygdala-frontal connectivity during emotional face processing from childhood into young adulthood. Hum Brain Mapp 37:1684–1695.

Xia M, Wang J, He Y (2013) BrainNet viewer: a network visualization tool for human brain connectomics. PLoS One 8:e68910.

Yan CG, Craddock RC, He Y, Milham MP (2013) Addressing head motion dependencies for small-world topologies in functional connectomics. Front Hum Neurosci 7:910.

Yarkoni T, Poldrack RA, Nichols TE, Van Essen DC, Wager TD (2011) Large-scale automated synthesis of human functional neuroimaging data. Nat Methods 8:665–670.

Yuan W, Altaye M, Ret J, Schmithorst V, Byars AW, Plante E, Holland SK (2009) Quantification of head motion in children during various fMRI language tasks. Hum Brain Mapp 30:1481–1489.