Face-Selective Responses Present in Multiple Regions of the Human Infant Brain

Heather L. Kosakowski[1]*, Michael A. Cohen[1,2], Lyneé Herrara[3], Isabel Nichoson[4], Nancy Kanwisher[1], Rebecca Saxe[1]

[1] Department of Brain and Cognitive Sciences, Massachusetts Institute of Technology, 77 Massachusetts Avenue, Cambridge, MA.

[2] Department of Psychology and Program in Neuroscience, Amherst College, 220 South Pleasant Street, Amherst, MA.

[3] Psychology Department, University of Denver, 2155 South Race Street, Denver, CO.

[4] Haskins Laboratory, Yale University, 300 George Street, New Haven, CT.

* Correspondence to: heatherlkosakowski@gmail.com

Abstract

Faces are a rich source of social information. How does the infant brain develop the ability to recognize faces and identify potential social partners? We collected functional magnetic neuroimaging (fMRI) data from 49 awake human infants (aged 2.5-9.7 months) while they watched movies of faces, bodies, objects, and scenes. Face-selective responses were observed not only in ventral temporal cortex (VTC) but also in superior temporal sulcus (STS), and medial prefrontal cortex (MPFC). Face responses were also observed (but not fully selective) in the amygdala and thalamus. We find no evidence that face-selective responses develop in visual perception regions (VTC) prior to higher order social perception (STS) or social evaluation (MPFC) regions. We suggest that face-selective responses may develop in parallel across multiple cortical regions. Infants' brains could thus simultaneously process faces both as a privileged category of visual images, and as potential social partners.

Introduction

From birth, infants see faces often [1–3], prefer to look at faces more than non-face visual stimuli [4–9], and readily distinguish a familiar face from an unfamiliar one [10,11]. Infants' brains visually process faces as a distinct object category [12] but it is not known when in development infants' brains begin to process the social attributes of faces. In adult brains, perception of faces not only evokes activity in face-selective visual regions [13–15] but also in higher-order brain regions that represent multi-modal social perception and evaluation [16–18]. This paper addresses the open, and hotly debated question, of how this suite of face-selective regions develop. We test two contrasting hypotheses: (1) face-selective responses develop sequentially, with visual perception regions responding selectively to faces earlier in development than higher-order social regions, or (2) face-selective responses emerge in such that perceptual and social regions come to represent the visual category and social meaning of faces simultaneously. Here we use functional magnetic resonance imaging (fMRI) data from a large group of awake human infants to test whether face-selective responses in social regions emerge later in development than visual perceptual regions.

Serial Hypothesis of Cortical Development

One prominent hypothesis, which we call the "Serial Hypothesis" of development, holds that distinct brain regions become face-selective in a feed-forward developmental sequence (Figure 1a; for recent reviews see: [19–21]). On this view, face-selective responses first emerge in subcortical regions as a response to the salient visual properties of faces, then in higher-level visual cortex as infants gain experience looking at faces, and finally in higher-order social cognition regions as infants learn the connection between faces and social value. By analogy to the conspecific detection mechanism in newly hatched chicks [22], the first brain region to respond distinctively to faces is likely to be subcortical. This subcortical region would contain a rough template of face-like images (e.g., high contrast shapes with an upper field bias), which would then drive infants to look at faces. Because subcortical regions are small and deep in the brain, they are very hard to measure in human or non-human primate infants, but plausible sites for an innate subcortical face template in primates include the superior colliculus, pulvinar, and/or amygdala [6].

Next, the frequent appearance of face-like images in infants' visual experience would drive bottom-up, self-organizing specialization of regions in the cortical visual system [20,23]. The Serial Hypothesis supposes that high-level visual regions initially self-organize along low-level features (e.g., retinotopy, spatial frequency, curvilinearity), called a proto-organization [20]. For example, the area that becomes the fusiform face area (FFA) would acquire a face-selective response profile

gradually, as infants gain visual exposure to the combination of curvy, low spatial frequency environmental input characteristic of faces [24]. In line with this hypothesis, adult face-selective regions respond more to images with curvilinear features than images with rectilinear features, even in the absence of faces [25,26] and macaques raised without exposure to faces do not have face-selective responses that can be measured with fMRI [27].

Finally, once high-level visual regions have formed a visual category of faces, even higher order brain regions would then be able to associate faces with multi-modal information about social meaning and emotional value. In adults, the social and emotional meaning of faces are encoded in (among other regions) superior temporal sulcus (STS) and medial prefrontal cortex (MFPC). The response profiles of these regions are distinct in adults from the perceptual face responses in FFA. The STS has been implicated in a variety of social perception tasks [17] including facial movement [28], social interaction perception [29], voice perception [30,31], and multimodal face and voice perception [32,33]. The MPFC is specifically involved in abstract social and emotional processes, such as social evaluation [34–36], self-referential processing [35,37–41], and emotion attribution [16,35,42–47].

The Serial Hypothesis of development predicts that responses to faces in STS and MPFC would develop later, potentially much later, than category selective regions in VTC [20–22]. One source of this prediction is the evidence that anatomical development proceeds from primary sensory areas to association cortices [48–53] (for review see: [54]). Signatures of cortical maturation including expansion [55], increased sulci depth [56], and myelination [57–59] occur in primary sensory areas earlier in development than association areas. Not only does anatomical development of the brain occur in series, but signatures of neural function such as glucose metabolism [60] and synaptogenesis [61] reach their peak in primary sensory areas prior to association areas. Perhaps most surprisingly, new neurons are still migrating and being integrated into PFC well into the second year of life [62] – a process that completes in primary sensory areas around the time of birth [63]. This evidence has led to the hypothesis that due to the structural immaturity of association cortex, functions associated with those regions in adults are not yet present in infants. Yet, it is equally plausible that rather than anatomical maturation preceding neural function, perhaps neural function and anatomical maturation are intrinsically reciprocal in brain development.

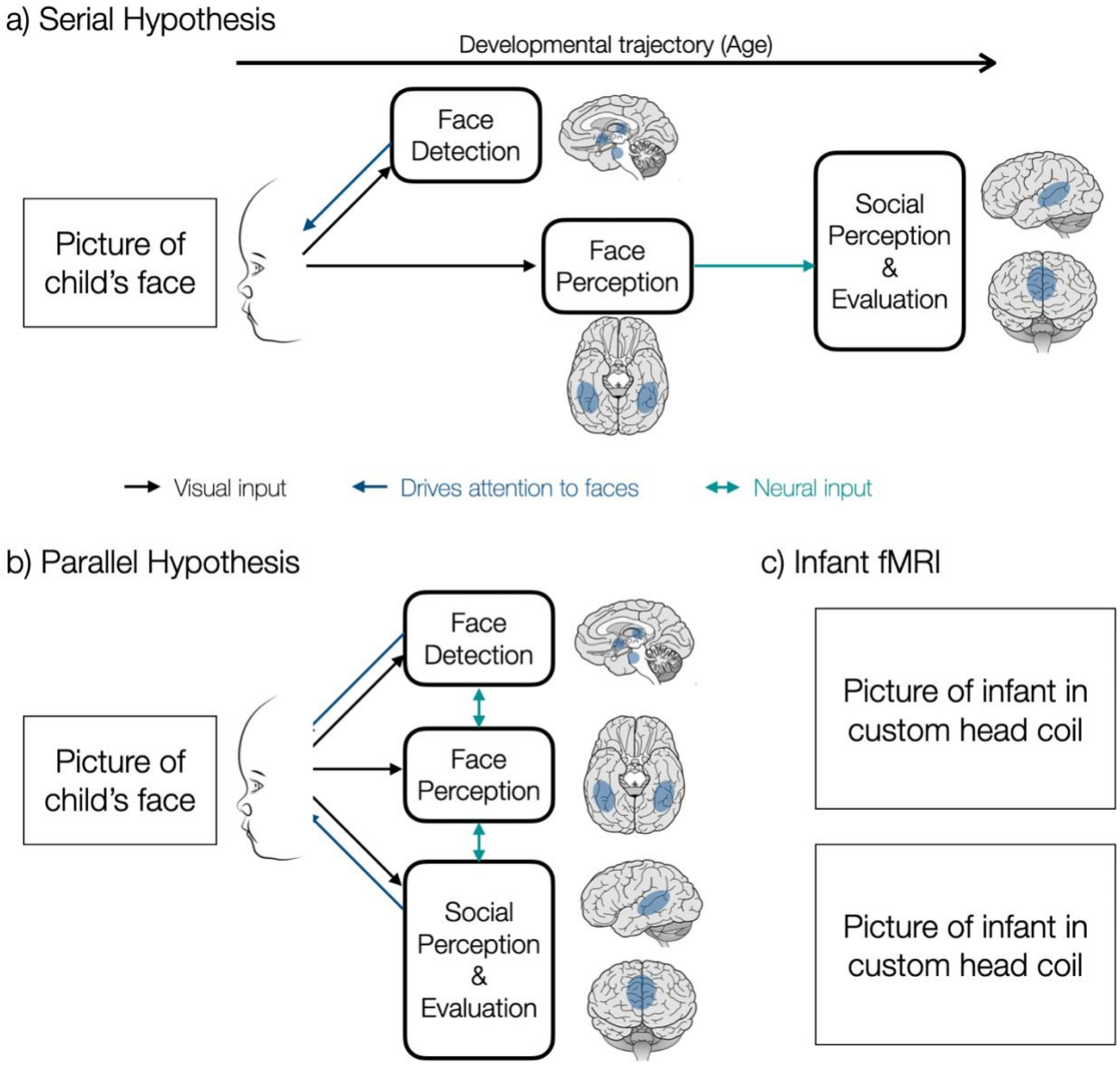


Figure 1. Alternative hypotheses make different predictions about when face-selective responses emerge in infants' brains. a) The Serial Hypothesis predicts that the earliest face-selective responses arise in an innate subcortical face-template, which drives infants to attend to face-like images. Visual experience with faces then causes self-organization of face-selective responses in ventral-temporal cortical regions, like FFA. Subsequently, higher-order regions like STS and MPFC can associate face shapes with emotional and social meaning. b) Alternatively, the Parallel Hypothesis proposes that from early in infancy, infants' attention to faces is driven by both perceptual and socio-emotional processes; and that face-selective responses in STS and MPFC arise simultaneously with, and potentially provide input to, perceptual representations of faces in ventral-temporal cortex. c) For each MRI visit, swaddled infants were placed in a custom 32-channel infant head coil. Movies were projected in a mirror over their eyes. Example face stimuli pictured (a-b). Blue areas were added to brain images created by Milan Vukelić and available for Public License at: https://www.behance.net/gallery/5399443/BRAINS-%28cognitive-neurosciences%29. Photo credit for images in c to Caitlin Cunningham Photography.

118 **Parallel Hypothesis of Cortical Development**

119 By contrast to the Serial Hypothesis of development, we propose a Parallel Hypothesis, in
120 which responses to the social meanings and emotional value of faces (by reverse inference in STS
121 and MPFC, respectively) emerge just as early in development as perceptual responses to faces in
122 FFA (Figure 1b). Substantial behavioral and neural evidence supports this hypothesis. Behaviorally,
123 infants' active attention to faces cannot be fully explained by a face-template or preference for face-
124 like images. Neonates and young infants do not merely prefer to look at faces over other stimuli, but
125 choose to look at some faces more than others, based on both visual (e.g., skin tone, sex, facial
126 expression) and non-visual (e.g., speaker language, prosody, social behavior) properties [11,64–72].
127 These behavioral data suggest that infants' brains represent not just the perceptual form, but also the
128 social meaning and emotional value of faces. Consistent with the predictions from behavioral data,
129 studies using functional near-infrared spectroscopy (fNIRS), electroencephalography (EEG), and
130 positron emission topography (PET) in human infants show responses to faces in the superior
131 temporal sulcus (STS) as early as a few days after birth [73–80] and in the medial prefrontal cortex
132 (MPFC) by 5-months (the earliest that has been reported) [74,77,80–82].

133 However, no previous study has directly tested whether face-selective responses in human
134 infants emerge in subcortical regions followed by visual higher-level visual cortex, followed by
135 social-emotional regions, as predicted by the Serial Hypothesis, or simultaneously across these
136 regions, as predicted by the Parallel Hypothesis. The main limitation is that responses in subcortical
137 and ventral visual regions cannot be measured using surface-based neuroimaging techniques like
138 fNIRS and EEG.

139 **Current study**

140 For the current study, we analyzed fMRI data from 49 infants (Figure 1c) while they watched
141 videos of faces, bodies, objects, scenes, and an abstract baseline. Infants ranged in age from 2 to 9
142 months, meaning that the oldest infants had three times as much visual experience as the youngest
143 ones. Thus, these data allow us to test distinctive predictions of the two hypotheses. The Serial
144 Hypothesis makes three predictions. First, infants as a group may not yet have any face-selective
145 responses in the purported latest developing regions (STS and MPFC). Second, face-selective
146 responses should be correlated with age, particularly in cortical regions. In visual perceptual regions
147 (e.g., FFA), face-selective responses should increase with infants' age and visual experience; and if
148 there is any face-selective response in STS or MPFC, it should be more selective in the older infants.

149 Third, a subset of infants (plausibly, the youngest ones) should have face-selective responses in FFA
150 but not yet in STS and MPFC.
151 Conversely, the Parallel Hypothesis makes a different set of predictions. If face-selective
152 responses appear early and simultaneously across multiple brain areas (Figure 1b), then infants as a
153 group should show face-selective responses in subcortical, visual perceptual, and social-emotional
154 regions. Second, face-selective responses should either be present in all infants from the earliest ages
155 or increase simultaneously in all regions. Third, any subset of infants who have face-selective
156 responses in FFA should also have face-selective responses in STS and MPFC.

157 # Results

158 **Face-selective responses observed in infant STS and MPFC**

159 One group of infants was scanned on an older infant coil [83] with a sinusoidal acquisition
160 sequence [84] (*Coil 2011;* n=26) and a second group of infants was scanned on a newer infant coil
161 [85] with a standard acquisition sequence (*Coil 2021*; n=23), resulting in higher SNR (see Methods).
162 For the first set of analyses, we treat the two datasets as independent tests of the hypothesis.
163 As the Serial and Parallel Hypotheses make contrasting predictions for whether infants as a
164 group will have face-selective responses in the STS and MPFC, we first look in these regions. In a
165 group random effects analysis, we observed responses to faces that were greater than the response
166 to the average of non-face conditions in bilateral STS and MPFC, in both the *Coil 2011* data and the
167 *Coil 2021* data (Figure 2).
168 To test whether face responses in STS and MPFC are significantly larger than responses to
169 each control condition (i.e., are "face selective" [15]), we used a functional region of interest (fROI)
170 analysis. We identified the top 5% of voxels that responded more to faces than the average of the
171 other categories within each broad region in individual infants and then extracted responses in those
172 voxels to all four conditions in independent, left-out data from the same infant (see Methods). In the
173 *Coil 2011* fROI analysis (Figure 2a, n=18), STS and MPFC face responses were significantly and
174 substantially greater than the response to each of the three other stimulus categories (all $Ps<0.02$;
175 Table S1). These results were replicated in the *Coil 2021* fROI analysis (n=20, Figure 2b; all
176 $Ps<0.003$; Table S1) and when data were combined across both coils (n=38, 35 individuals; Figure
177 3a-b; all $Ps<0.00008$; Table S1).

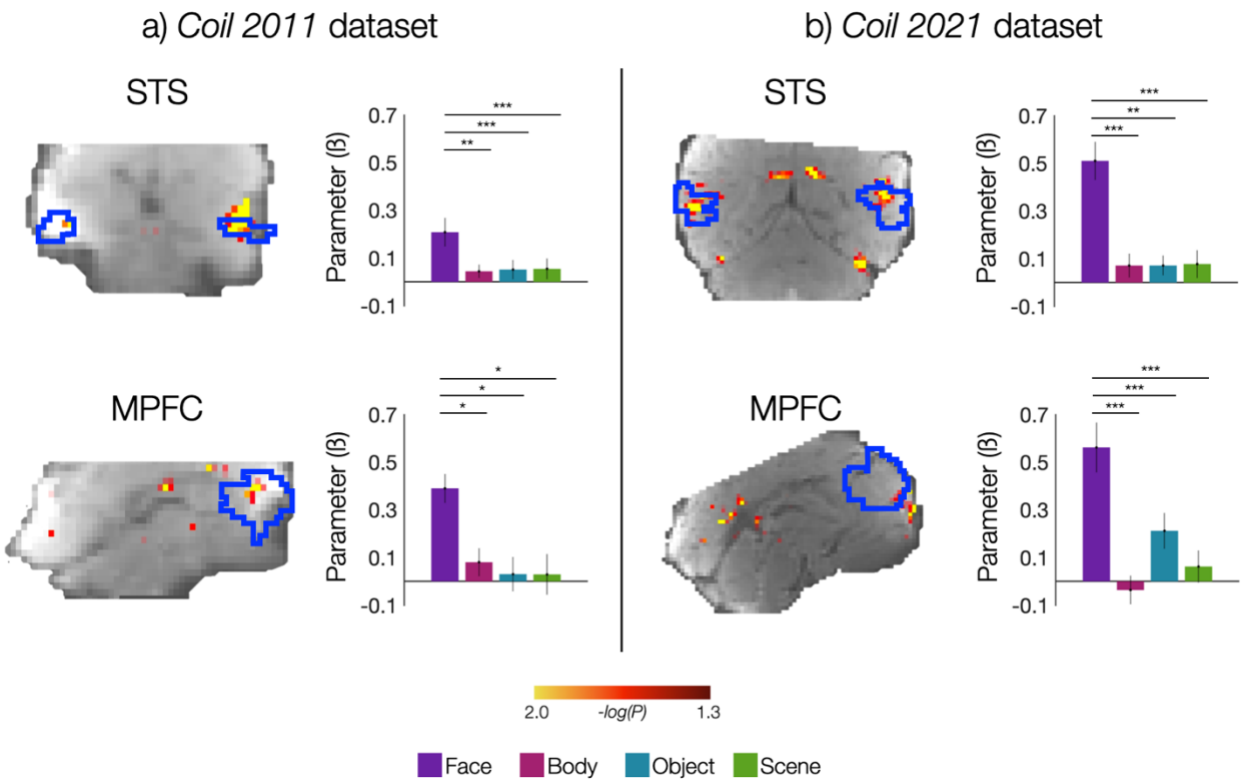


Figure 2. Face responses in infant STS and MPFC are selective. Group whole brain random effects analysis of *Coil 2011* data (a; n=26) and *Coil 2021* data (b; n=23) revealed face responses (faces>(bodies+objects+scenes)) in both STS (top) and MPFC (bottom). Note that the precise anatomical location of STS and MPFC activations between *Coil 2011* and *Coil 2021* datasets cannot be directly compared due to differences in distortion and resolution (Figure S2). Group maps are displayed on representative functional images from each coil; threshold $P<0.05$, uncorrected, for visualization purposes. Cortical area outlined blue represents the anatomical search space. An fROI analysis of *Coil 2011* (n=20, 18 individuals) revealed face responses (purple) that were significantly greater in independent data than the response to bodies (pink) objects (teal) and scenes (green) in STS and MPFC. These results were replicated in the *Coil 2021* dataset (b). Symbols: $^{*}P<0.05$; $^{**}P<0.01$; $^{***}P<0.001$. Error bars indicate within-subject SE [86].

With the *Coil 2011* and *Coil 2021* datasets combined, we next conducted the same fROI analyses varying the number of voxels selected from the top 1% of voxels to 100% of voxels in each of the parcels, and tested responses in held out data. In both STS and MPFC, no matter how large the subset of voxels selected, the observed face response was significantly greater than the response to each non-face condition (Figure 3a-b) indicating face-selective responses in infant STS and MPFC are robust and do not depend on the threshold for fROI definition.

Thus, infants as a group have face-selective responses in STS and MPFC, regions that represent social meaning and emotional value in adults. The presence of face-selective responses in

infant STS and MPFC importantly constrains the Serial Hypothesis: if face selectivity develops serially, the process must occur within the first few months of life.

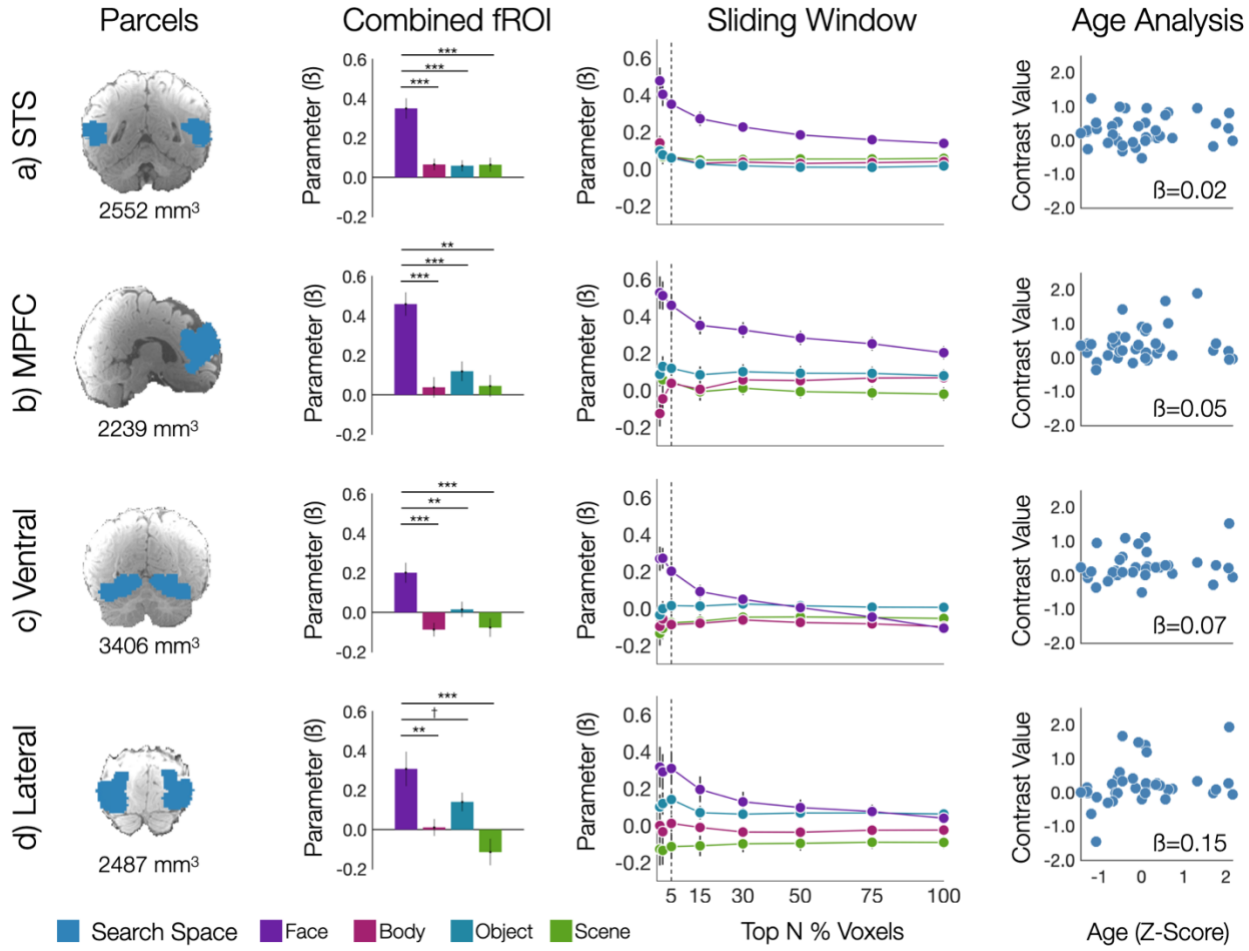


Figure 3. Cortical face-selective responses are not correlated with age. In fROI analyses collapsed across *Coil 2011* and *Coil 2021* datasets (n=38, 35 individuals), we used four large anatomical search spaces (first column, blue, projected onto an infant anatomical image) of STS (a), MPFC (b), ventral-temporal cortex (c), and lateral-occipital cortex (d). Bar charts show the average response across participants in each fROI (top 5% of voxels) to each stimulus category (compared to baseline) in data independent of that used to define the fROI. Error bars indicate within-subject SE [86]. Symbols used to report one-tailed statistics from linear mixed effects models: $^{\dagger}P<0.1$; $^{*}P<0.05$; $^{**}P<0.01$; $^{***}P<0.001$. Line graphs show selectivity analysis with different proportions of voxels selected with the average response in independent data for faces (purple), bodies (pink), objects (teal), and scenes (green). The vertical dashed line marks the top 5% and corresponds to the bar charts. Error bars are standard error of the mean. The scatter plots show that the contrast value (the difference between the face response and the average response to all non-face conditions) for each fROI was not correlated with age (all *Ps*>0.19).

Infants' cortical responses to faces were not correlated with age

The two hypotheses make contrasting predictions for whether face-selective responses in cortical regions should increase with age and visual experience. We tested these predictions in two visual regions (FFA and OFA) and in two social-emotional regions (STS and MPFC). To maximize the potential that we would be able to detect changes with age and experience, we combined the *Coil 2011* and *Coil 2021* datasets. Importantly, the amount of low-motion data included from each infant was not correlated with age (Fig. S1).

First, we confirmed that there is a selective response to faces in the visual regions of these infants (see [12], for a complementary analysis of *Coil 2021* dataset alone in FFA- and OFA-specific parcels). To account for differences in distortion between *Coil 2011* and *Coil 2021* datasets (See Fig. S2 for example images), we used an expanded parcel of the ventral temporal cortex which included the location of the FFA (ventral parcel) and an expanded parcel of lateral occipital cortex that included the location of the OFA (lateral parcel; see Methods). In an fROI analysis (n=38 datasets from 35 infants), the face response in the ventral parcel (Figure 3c; Table S1) was substantially and significantly larger than the response to each of the non-face conditions (all $Ps<0.005$). In the lateral parcel (Figure 3d; Table S1), we found weak evidence for a face-selective response (all $Ps<0.06$).

Next, we checked whether the difference between responses to face and non-face stimuli increased with age in any cortical region. In a linear mixed effects (LME) analysis with age as a predictor and contrast value (i.e., the response to faces minus the average response to non-faces conditions) as the dependent variable (Figure 3c-d, last column), infants' age did not predict the contrast value in either of the visual regions (LME, all $Ps>0.2$; Table S2). Similarly, there was no effect of age on the contrast value in either STS or MPFC parcels (all $Ps>0.5$; Table S2).

Thus, we did not find evidence that face-selective responses in cortical regions increase with age. Instead, face-selective responses in both visual and social-emotional regions are relatively constant between ages 2 and 9 months in all regions tested.

No subset of infants showed face-selective responses in subcortical and visual regions but not social-emotional regions.

The Serial Hypothesis distinctively predicts that a subset of infants (plausibly, the youngest ones) should have face-selective responses in subcortical regions and FFA, but not yet in STS and MPFC. By contrast, the Parallel Hypothesis predicts that any subset of infants who have face-selective responses in FFA should also have face-selective responses in STS and MPFC. We tested these predictions in four complementary ways.

First, we considered the full sample collected with *Coil 2021*, because these data had the fewest distortions and thus provided the best measurements in ventral visual and subcortical regions. Thus, in these infants we tested whether subcortical face-responses emerge earlier than ventral visual regions, and whether ventral visual face responses emerge earlier than STS or MPFC. Group random effects of the *Coil 2021* dataset did not reveal subcortical responses to faces when compared to the average of the other conditions (see Fig S3 for whole brain analyses of *Coil 2011* data); however fROI analyses are more sensitive.

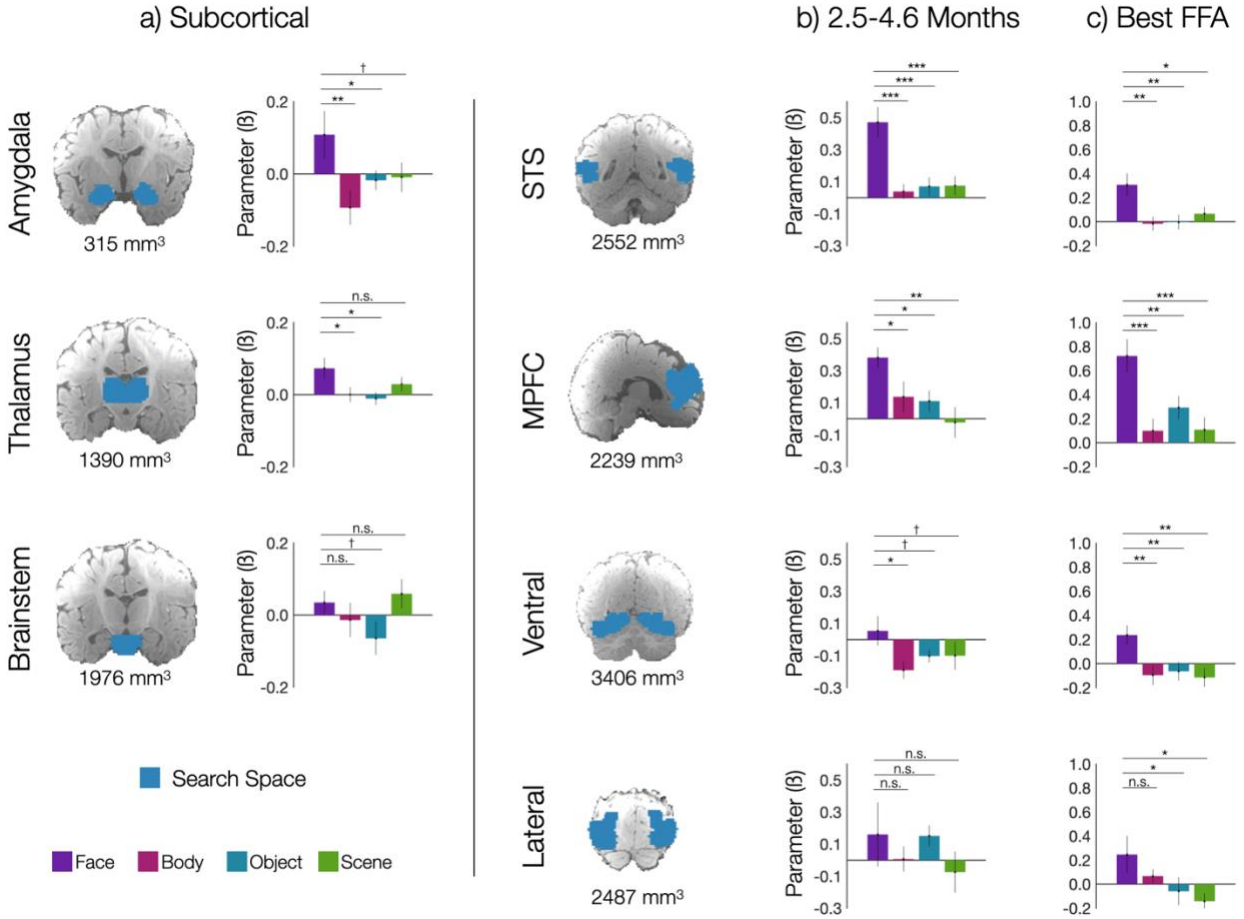


Figure 4. All subsets of infants have face-selective responses in STS and MPFC. (a) An fROI analysis of subcortical areas in infants measured in the *Coil 2021* dataset revealed that the face response (purple) in the amygdala (a, top) and thalamus (a, middle) was significantly greater than the response to bodies (pink) and objects (teal) but not statistically different than the response to scenes (green). In the brainstem of these infants (a, bottom), no face-selective response was detected. (b) In an fROI analysis of all infants younger than 5.0 months, both STS and MPFC had face-selective responses, the ventral parcel had a marginally face-selective response, and the lateral parcel had no face-selective response. (c) In the 10 infants with the greatest contrast value (response to faces minus the average response to non-face conditions) in the FFA, both STS and MPFC had face selective regions, while the lateral parcel had a region that responds significantly more to faces than objects or scenes, but not bodies. Error bars indicate within-subject SE [86]. Symbols used to report

one-tailed statistics from linear mixed effects models: $^{n.s.}P>0.1$; $^{\dagger}P<0.1$; $^{*}P<0.05$; $^{**}P<0.01$; $^{***}P<0.001$.

We defined parcels to search for subcortical responses in superior colliculus, amygdala, and thalamus. The superior colliculus, a structure found in the brainstem, is thought to increase the salience of face-like images and thus drive infants' face looking behavior [20,22]. To search for the (very small) superior colliculus, we used our fROI method within an anatomically defined parcel covering the brainstem (Figure 4a; Table S1). While the brainstem contains a region responding numerically more to faces than to objects (P=0.06), it was not statistically different than the response to either bodies (P=0.2) or scenes (P=0.6). The amygdala and thalamus are two other plausible sites for the putative face template. An fROI analysis (Figure 4a; Table S1) revealed weak evidence for face-selective responses in the amygdala (all $Ps\leq0.06$) and the thalamus ($Ps\leq0.11$). In the same group of infants (*Coil 2021* data), we observed face-selective responses in the ventral (all $Ps<0.001$) and lateral parcels (all $Ps<0.03$) as well as STS and MPFC parcels (reported above; Figure 2). These differences between regions were not due to differences in temporal signal-to-noise ratio (tSNR): subcortical, STS, and ventral parcels had a higher tSNR than MPFC and lateral parcels (Figure S4). In the *Coil 2021* dataset alone, age was not a good predictor of contrast magnitude in any of the cortical parcels (MPFC ß=0.03, P=0.9; STS ß=-0.01, P=0.9; Ventral ß=0.12, P=0.3; Lateral ß=0.21, P=0.3) or the thalamus (ß=0.04, P=0.8) but was a marginal predictor of contrast magnitude in the brainstem (ß=0.07, P=0.09) and amygdala (ß=0.12, P=0.1). Thus, we find no evidence that face-selective responses emerge earlier in subcortical than in cortical regions, or earlier in visual perception regions than in social-emotional regions, in the *Coil 2021* dataset.

Second, we considered just the youngest infants in our combined sample, those less than 5.0 months old. In an fROI analysis of these young infants (n=15, 2.5-4.6 months, mean=3.6 months; Figure 4b; Table S3), we found weak evidence for a face-selective response in the ventral parcel (all $Ps<0.09$) but not the lateral parcel (all $Ps>0.1$). Yet both the STS and MPFC also exhibited face responses that were substantially and significantly greater than the response to each of the non-face conditions (all $Ps\leq0.03$). Taken together, there was no evidence that visual perception regions exhibit more robust face-selective responses than social-emotional regions in the youngest infants.

Third, we examined the infants with the most face-selective response in FFA. We selected the 10 infants with the largest contrast value (the response to faces subtracting the average response to non-face conditions) in the ventral fROI. In these infants (Figure 4c; Table S3), we also found face-selective responses in both the STS (all $Ps<0.02$) and the MPFC (all $Ps<0.009$).

Fourth, in a final effort to determine whether the FFA develops before the STS and/or MPFC, we randomly sampled a subset of 15 infants and tested whether FFA, STS and/or MPFC showed a face-selective response. Of 2500 random sub-samples (Figure S5), all three fROIs had face-selective responses in 36.8% (919 instances) of subsamples while 0.9% (23 instances) of subsets showed no face-selective response in any region. In 1.2% of subsets (31 instances), infants had face-selective responses in VTC but neither STS or MPFC, whereas face-selective responses in STS and/or MPFC but not VTC occurred in 48.5% (1212 instances) of subsets. Thus, subsets of infants who have face-selective responses in VTC but not STS or MPFC are very rare and non-representative. Taken together, these results indicate that face-selective responses in visual-perceptual regions do not emerge prior to face-selective responses in higher-order social-emotional regions.

Discussion

Here we tested two alternative hypotheses about the development of face-selective responses in human cortex. Contrary to the predictions of the Serial Hypothesis, we found no evidence that face selectivity emerges sequentially in subcortical regions, followed by ventral visual regions, followed by "higher order" social and emotional regions. Specifically, we found evidence for face-selective responses in STS and MPFC across our dataset and these responses were not correlated with age. Most tellingly, we did not observe any evidence for a subset of infants with face-selective responses in FFA but not in STS and MPFC. These results, however, are consistent with the Parallel Hypothesis, which proposes that face-selective responses emerge simultaneously across perceptual and social-emotional regions.

We note several limitations of the current study. Although we failed to find a correlation between infants' age and the magnitude of face-selective responses in any region, it is possible that a true correlation was masked by the variability of data quality across infants. Awake infants are a very challenging population for fMRI scanning. Using dynamic stimuli and custom-built infant-sized head coils [83,85], we acquired fMRI data from a large group of young infants [12], but inevitably, each infant yielded a different amount of usable data. Because the amount of data per infant was low (though not correlated with age), the estimate of face selectivity in each individual infant is noisy. Also, the age range we tested constituted a three-fold increase in visual experience but was still small relative to the human lifespan. The youngest infants who successfully participated were already 2 months old, and the oldest infants were only 9 months old. So, face-selective responses in cortical regions may increase with age prior to age 2 months, or after 9 months, or with a very gradual slope

334 that requires more data to detect. For these reasons, our failure to find a correlation between age
335 and face selectivity is only weak evidence against the Serial Hypothesis.
336 The stronger evidence against the Serial Hypothesis was the finding that no subset of infants
337 in this population had face-selective responses in FFA but not STS or MPFC. One possible
338 limitation of this analysis is that the temporal signal-to-noise ratio (tSNR) and spatial distortions of
339 the images were not homogeneous across the brain. For example, the subcortical regions, which are
340 anatomically smaller, were much more distorted in the *Coil 2011* dataset than the *Coil 2021* dataset.
341 The hint of face-selective responses in the amygdala and thalamus in the *Coil 2021* dataset is
342 intriguing but should be confirmed in future studies. We did not have enough data to test whether
343 subcortical regions had face-selective responses in the youngest infants, as predicted by the Serial
344 Hypothesis. On the other hand, consistent with the Parallel Hypothesis, nearly every subset of
345 infants we checked, including just the youngest infants (2.5-4.6 months), had face-selective
346 responses in STS and MPFC.
347 **Potential Functions of STS and MPFC in Infancy**
348 What functions might STS and MPFC be serving in young infants? For simplicity of
349 argument, we have up to this point treated STS and MPFC as part of a single class of "higher order"
350 regions that respond to social-emotional face features. However, these regions are anatomically
351 distant from each other and associated with different functions in adults. The STS is implicated in
352 social perception, specifically in representing the meaning of facial movements. The MPFC is
353 implicated in social evaluation, representing the value of a social stimulus or event with respect to
354 oneself or others. From behavioral evidence, it is plausible that infants engage in both kinds of social
355 cognition; and from (limited) fNIRS evidence, it is plausible that the STS and MPFC serve
356 analogous functions in infants to those they serve in adults.
357 For example, in adults a region of STS responds more to dynamic faces and voices than any
358 other stimulus, including still faces [17,87]. Infants prefer to look at dynamic faces more than still
359 faces [88,89], and in studies using fNIRS, channels near STS in infants responds more to dynamic
360 faces compared to dynamic objects or scenes [74,76,90,91]. Another region of STS in adults
361 responds to cues of animacy and social interaction in amorphous shapes or point-light display [29].
362 Infants use the same cues to identify agents and interpret actions in behavioral experiments [92–96],
363 and an fNIRS channel near STS responded to point light displays of face movement compared to
364 scrambled motion [75,97]. Thus, it is plausible that in infants and adults, regions of STS are
365 responsible for encoding other people's facial and bodily movements as meaningful actions.

By contrast, in adults MPFC plays a distinct role from the STS. Observing social relationships, inferring others' emotions, and engaging in self-relevant tasks activate areas of MPFC [33–46]. Substantial behavioral evidence indicates that infants evaluate social stimuli in terms of their valence and self-relevance. Infants prefer to look at smiling faces and prosocial agents more than neutral, negative, or anti-social faces and agents [68,94,98,99] and, self-relevant faces such as those that match the race or gender of an infant's primary caregiver, garner longer looking times [67,69,100–104]. Infants also have a preference for people that sing the same song [105] or speak the same language [71] as their primary caregiver. These behavioral capacities may be at least in part supported by MPFC. FNIRS channels over infant MPFC respond more to smiling faces compared to neutral faces [106,107], and to people that appear to engage directly with the infant by using a direct gaze [81,82], playing hand games [79], or speaking in an infant-directed manner [108]. Taken together, these results provide evidence that infant MPFC, like adult MPFC, preferentially responds to valued, emotional, and self-relevant social stimuli.

Process of Parallel Cortical Development

In summary, we argue that face-selective responses may emerge in parallel during infancy in both perceptual and higher-order association cortices. In adults, these brain regions have functions that range from detection and recognition of facial identity (e.g., FFA) to perception of facial movement and expression (STS) to the extracting the emotional meaning and value of the social interaction (MPFC). How could face selectivity emerge in parallel across all of these regions? One possibility is that even prior to maturation of sensory areas, social interactions elicit a reward response which drives young infants to attend to pro-social, self-relevant aspects of their environment. Under this framework, seeing engaging human faces, hearing self-relevant voices, and pro-social touch would activate infants' reward network, resulting in increased attention to socially relevant stimuli. In support of this hypothesis, infant STS responds to social perception in the auditory [79,91,109,110], visual [74,76,90,91], and tactile [111] domains while MPFC is part of the reward network [112–115] and responds to socially self-relevant environmental cues in infants [79,81,82,106–108]. In infants, affective touch enhances infants' attention to faces [116] and increases functional coupling between STS and MPFC [117], while auditory language modulates activity in STS and MPFC [110,118]. Social perception and interaction through touch and audition begins earlier and is more sustained in neonates' early experiences than visual input, in part because of neonate's limited visual acuity. Thus, social experience in non-visual modalities could prime and potentiate the receptivity of STS and MPFC to faces, particularly those associated with infant-

398 directed speech or touch. In this way, social reward could drive infants' preferences to look at
399 familiar, friendly, or socially valued people, shaping the input received by visual face-perception
400 regions. On this view, during infancy, face perception in FFA, social perception in STS and social
401 evaluation in MPFC may have reciprocal and mutually reinforcing connections. The consequence of
402 such connections would be simultaneous, parallel development of face-selectivity in all of these
403 cortical regions.

404 The finding that MPFC has functionally specific responses in infancy also has implications
405 for broader theories of cortical development. By anatomical metrics like myelination and cortical
406 thinning, prefrontal cortices are the slowest cortical areas to mature. Anatomical development of
407 PFC is implicated in late-developing cognitive functions such as executive function and language
408 [48–54,119–121]. Thus, many researchers have assumed the prefrontal cortex is not a major player
409 in cognitive function during the earliest stages of development. In contrast, our results are consistent
410 with mounting evidence that infant PFC has remarkably sophisticated function despite structural
411 immaturity [122]. Prior studies have reported activity in lateral PFC that corresponds to adult
412 functional organization. For example, a region in lateral PFC in infants responds more to sequences
413 with statistical regularity compared to unstructured input [123–125]. Another area in infant PFC
414 responds more to native language compared to foreign language or other non-speech sounds [126–
415 130] and putative language areas have highly specific connection with orthographic-specific
416 perceptual regions in VTC [131]. Thus, Werchan and Amso [132] have argued that PFC function is
417 not immature in infants, but functions adaptively for infants' age-specific ecological niche; Deheane-
418 Lambertz and Spelke [133] argued for PFC as key player in infant cognitive development; and Saygin
419 and colleagues [131,134] have argued that long-range structurally connectivity with PFC constrains
420 functional development. Our current results provide further evidence in favor of such views,
421 incorporating the social functions of MPFC as well.

422 **Summary**

423 In sum, using fMRI in awake human infants, we find that in addition to visual perception
424 regions, the STS and MPFC also have face-selective responses early in infancy. These results suggest
425 that human infant brains not only perceive faces as a special class of objects, but also process faces
426 as self-relevant potential social partners. Our findings about the origins of face selectivity accord
427 with a broader theoretical view about the origins of the mind: infants are not passive statisticians
428 analyzing data in a purely data-driven fashion [135], but actively participate in the construction of
429 their own minds, choosing which input to process and hypotheses to test in pursuit of their own

agenda. The social value of faces may thus be reflected in the order and timing of development of face-selective responses across multiple regions of the human infant brain.

Methods

Subjects

Infants were recruited from the Boston metro area through word-of-mouth, fliers, and social media. Parents of participants were provided parking or, when travel was a constraint for participation, reimbursed travel expenses. Participants received a small compensation for each visit and, whenever possible, printed images of their brain. We recruited 86 infants (2.1-11.9 months; mean age = 5.2 months; 41 female; 42 *Coil 2021*) and recovered usable data (see data selection) from 49 infants (*Coil 2011*, n=26, 2.5-8.7 months, 13 female; *Coil 2021*, n=23, 2.1-9.7 months, 11 female). These data were previously used to test if infants have category-selective responses to faces, bodies and scenes in ventral temporal cortex. Results from that investigation are reported in [12].

Stimuli

Paradigm 1. Infants watched videos of faces, bodies, objects, and scenes [87], and a colorful, curvy, abstract baseline was used to maintain infants' attention. Videos were selected to be categorically homogeneous within blocks and heterogeneous between blocks. Each block was 18s and was composed of six 3s videos from the same category. Face videos showed one child's face on a black background. Object videos show toys moving. Body videos showed children's hands or feet on a black background. Scene videos showed natural environments. Baseline blocks were also 18s and consisted of six 3s videos that featured abstract color scenes such as liquid bubbles or tie-dyed patterns. Block order was pseudo-random such that all blocks played once prior to playing again. Videos played continuously for as long as the infant was content, paying attention, and awake.

Paradigm 2. Infants watched videos from the same five categories (faces, objects, bodies, scenes, curvy baseline) as in Paradigm 1. However, the videos were shortened to 2.7s, interleaved with still images from the same category (but not drawn from the videos) presented for 300 ms. All blocks were 18s and included 6 videos and 6 images. Video and image order were randomized within blocks and block order was pseudorandom by category. Paradigm 2 contained one additional block depicting hand-object interactions (not included in the present analysis).

Data Collection

Infants were swaddled if possible. A parent or researcher went into the scanner with the infant while a second adult stood outside the bore of the scanner. Infants heard lullabies (https://store.jammyjams.net/products/pop-goes-lullaby-10) for the duration of the scan. For data collected with *Coil 2011*, lullabies were played over a loudspeaker into the scanning room. For data collected with *Coil 2021*, lullabies were played through custom infant headphones (Figure 1c).

464 *Coil 2011*. For data collected with *Coil 2011* we used similar methods as previously reported [77]. We
465 used a custom 32-channel infant coil designed for a Siemens Trio 3T scanner [83]. We used a quiet
466 EPI with sinusoidal trajectory [84] with 22 near-axial slices (repetition time, TR=3s, echo time,
467 TE=43 ms, flip angle=90°, field of view, FOV=192 mm, matrix=64x64, slice thickness=3 mm, slice
468 gap=0.6 mm). The sinusoidal acquisition sequence caused substantial distortions in the functional
469 images (Figure S2a).
470 *Coil 2021*. Infants wore custom infant MR-safe headphones (Figure 1a). Infant headphones
471 attenuated scanner noises and allowed infants to hear the lullabies. An adjustable coil design
472 increased infant comfort and accommodated headphones as well as a variety of head sizes (Figure
473 1c). The new infant coil and infant headphones designed for a Siemens Prisma 3T scanner enabled
474 the use of an EPI with standard trajectory with 44 near-axial slices (repetition time, TR=3s, echo
475 time, TE=30ms, flip angle=90°, field of view, FOV=160 mm, matrix=80x80, slice thickness=2 mm,
476 slice gap=0 mm). Functional data collected with *Coil 2021* were less distorted (Figure S2b) than data
477 collected with *Coil 2011* (Figure S2a).
478 **Data Selection (subrun creation)**
479 To be included in the analysis, data had to meet criteria for low head motion. Motion criteria are the
480 same as our two previous studies [12,77]. Data were cleaved between consecutive timepoints having
481 more than 2 degrees or mm of motion, creating *subruns*, which contained at least 24 consecutive low-
482 motion volumes. All volumes included in a subrun were extracted from the original run data and
483 combined to create a new Nifti file for each subrun. Paradigm files were similarly updated for each
484 subrun. Volumes with greater than 0.5 degrees or mm of motion between volumes were removed
485 from all analyses. Total data collected and total data included in each analysis from each subject are
486 reported in [12].
487 Data collected within a 30-day window from a single subject were analyzed as one data
488 point. Subjects had to have at least 5 minutes of low-motion data to be included in the whole brain
489 analysis. For *Coil 2011*, this procedure resulted in 566.0 minutes of usable data (mean=5.9 minutes,
490 s.d.=11.3) from 27 subjects (mean age=4.7 months, s.d.=1.4; n=10 Paradigm 2). For *Coil 2021*, this
491 procedure resulted in 514.3 minutes of usable data (mean=6.5 minutes, s.d.=12.6) from 23 subjects
492 (mean age=5.7 months, s.d.=2.1; all Paradigm 2). Overall, 49 unique individual infants' data were
493 included in whole brain analyses (one subject contributed data to both *Coil 2011* and *Coil 2021*
494 datasets).

To be included in the fROI analysis, subjects had to have at least two subruns with at least 96 volumes in each subrun (one to choose voxels and the other to independently extract response magnitudes from the selected voxels). We included subjects in the fROI analysis for whom we had two timepoints and accounted for individual differences by using a subject random effects regressor and an age regressor in our linear mixed effects analyses (see fROI methods). This resulted in a final fROI dataset of 39 total data points from 36 unique individuals (2.5-9.7 months).

Preprocessing

Each subrun was processed individually. First, an individual functional image was extracted from the middle of the subrun to be used for registering the subruns to one another for further analysis. Then, each subrun was motion corrected using FSL MCFLIRT. If more than 3 consecutive images had more than 0.5 mm or 0.5 degrees of motion, there had to be at least 7 consecutive low-motion volumes following the last high-motion volume for those volumes to be included in the analysis. Additionally, each subrun had to have at least 24 volumes after accounting for motion and sleep TRs. Functional data were skull-stripped (FSL BET2), intensity normalized, and spatially smoothed with a 3mm FWHM Gaussian kernel (FSL SUSAN).

Data registration

All subruns were registered within subjects and then each subject was registered to a standard template. First, the middle image of each subrun was extracted and used as an example image for registration. If the middle image was corrupted by motion or distortion, a better image was selected to be the example image. The example image from the middle subrun of the first visit was used as the target image. All other subruns from each subject were registered to that subject's target image using FSL FLIRT. The target image for each subject was then registered to a template image using FSL FLIRT. For data collected with *Coil 2011*, the template image was the same one used previously [77]. For data collected with *Coil 2021*, a template image was selected from a single subject from whom we had a high-resolution anatomical image on which to display functional data. Given the distortion of the images and the lack of an anatomical image for each subject, traditional registration tools do not effectively register infant data. As such, we attempted to register each image using a rigid, an affine, and a partial affine registration with FSL FLIRT. The best image registration was selected by eye from the three options and manually tuned using the Freesurfer GUI for the best possible data alignment. Each image took between 2 and 8 hours of human labor to register. Images collected with *Coil 2021* were transformed to the anatomical space of the template image for visualization in Figure S2.

527 **Subject-level Beta and Contrast Maps**

528 Functional data were analyzed with a whole-brain voxel-wise general linear model (GLM) using
529 custom MATLAB scripts. The GLM included 4 condition regressors (faces, scenes, bodies, and
530 objects), 6 motion regressors, a linear trend regressor, and 5 PCA noise regressors PCA noise
531 regressors analogous to GLMDenoise [136]. Condition regressors were defined as a boxcar function
532 for the duration of each condition block (18s). Infant inattention or sleep was accounted for using a
533 single impulse nuisance ('sleep') regressor. The sleep regressor was defined as a boxcar function with
534 a 1 for each TR the infant was not looking at the stimuli, and the corresponding TR was set to 0 for
535 all condition regressors. Boxcar condition and sleep regressors were convolved with an infant
536 hemodynamic response function (HRF) that is characterized by a longer time to peak and deeper
537 undershoot compared to the standard adult HRF [137]. Next, data and all regressors except PCA
538 noise regressors were concatenated across subruns. PCA noise regressors were computed across
539 concatenated data and beta values were computed for each condition in a whole-brain voxel-wise
540 GLM. Subject-level contrast maps were computed as the difference between the face beta and the
541 average of all non-face (bodies, objects, and scenes) betas for each voxel using in-house MATLAB
542 code.

543 **Group Random Effects Analysis**

544 To test whether there was systematic overlap between areas of activation across subjects, we
545 conducted one group random effects analysis for eligible data collected with each coil. First, subject-
546 level contrast difference (faces – non-face) maps (see data inclusion and subject contrasts above)
547 were transformed to coil-specific template space (see data registration method above). Group
548 random effects analyses were performed using Freesurfer mri_concat and Freesurfer mri_glmfit.
549 Face responses were only considered if they were significantly greater than non-faces and fell within
550 group-constrained parcels (see next section for parcel definition).

551 **Parcels / Search Spaces.** Due to the distorted nature of the *Coil 2011* dataset, we opted to used
552 larger parcels than the standard FFA parcel we used previously [12] in order to include all available
553 data. We created lateral, ventral, STS, MPFC, and subcortical parcels using the Glasser atlas [138].
554 The lateral parcel included Glasser areas LO1, LO2, LO3, V4, V4t and PIT while the ventral parcel
555 included Glasser areas VMV1, VMV2, VMV3, VVC, PHA1, PHA2, PHA3, and FFC. For the STS
556 parcel, we used Glasser areas STSvp, STSva, STSdp, STSda, STV, and for the MPFC parcel we used
557 Glasser areas p24, d32, 9m, and p32. We also used the anatomically defined thalamus, brainstem,
558 and amygdala from MNI space.

Functional Region of Interest (fROI)

To determine if cortical responses are face-specific, we utilized a functional region of interest (fROI) approach. Due to the variable amount of data in each subrun for each subject and the impact this could have on reliable parameter estimates from the GLM, we first combined or split subruns to approximately equate the amount of data across subruns within subjects. For example, if a subject had three subruns and the first was 30 volumes, the second was 75 volumes and third was 325 volumes, then we concatenated the first two subruns to create one subrun and we split the third subrun into 3 resulting in a total of four subruns with approximately 100 volumes per subrun.

To constrain search areas for voxel selection, we used anatomically defined parcels (see parcels/search spaces) transformed to subject native space. We used an iterative leave-one-subrun-out procedure such that data were concatenated across all subruns except one prior to the whole-brain voxel-wise GLM and contrast were computed (described above). The top 5% most significant voxels for the contrast faces>(bodies+objects+scenes) within an anatomical constraint parcel were selected as the fROI for that subject, and the parameter estimates were extracted from a GLM on the left-out subrun. For all bar plots, beta values were averaged across participants and experiments.

To determine whether a region's response was category-selective, we fit the beta-values using a linear mixed effects model. In each region, we had an a priori hypothesis faces would elicit the largest response. So, in each model, we dummy-coded the other three control conditions, to test the hypotheses that the response to each control condition was lower than to the predicted face. Specifically, we fit a model in MATLAB with the expression:

```
fitlme(data ~ f1+f2+f3+age+sex+motion+(1|subject)+(1|coil))
```

where the three dummy-coded condition regressors are f1 (bodies), f2 (objects) and f3 (scenes). Fixed effects parameters of no interest were age, motion, and sex. Motion was the fraction of scrubbed volumes. Subject and coil were each coded as random intercepts for all models. In the *Coil 2011* analyses, we also included a fixed effect paradigm parameter ("para") because two slightly different versions of the experimental paradigm were used during data collection.

The response in a parcel was deemed selective if the fixed effect coefficient for each of the three control conditions was significantly negative, using a t-test. Because predictions are unidirectional, reported p-values are one-tailed. For example, a face parcel was only deemed face-selective if the parameter estimates for the body, object, and scene regressors were significantly negative.

Effects of infant age.

Effect of age on contrast magnitude: One metric of face selectivity is the difference between the response to faces and all other categories (i.e., the contrast value). To determine if the contrast difference between faces and all non-face conditions changed as a function of infant age, we used a linear mixed effects model for each parcel. Contrast values (i.e., face beta – average beta for all non-face conditions) was the dependent variable. The predictor of interest was age. Motion and sex were fixed effects of non-interest. Coil and subject were coded as random effects.

Acknowledgments: This research was carried out at the Athinoula A. Martinos Imaging Center at the McGovern Institute for Brain Research at MIT. The authors thank Nayanika Das and Somaia Saba for help with registrations; Steven Shannon, Atsushi Takahashi, and Boris Keil for technical support; members of Saxe Lab, and members of the Kanwisher Lab for help during recruitment and data collection; the Cambridge Writing Group, and members of the Saxe Lab and Kanwisher Lab for helpful comments on the manuscript; Michelle Hung and Kirsten Lydic for code review; Hannah LeBlanc for all the things; and all the infants and their families. **Funding:** We gratefully acknowledge support of this project by a National Science Foundation (graduate fellowship to HLK; Collaborative Research Award #1829470 to MAC), NIH (#1F99NS124175 to HLK; #R21-HD090346-02 to RS; #DP1HD091947 to NK; shared instrumentation grant S10OD021569 for the MRI scanner), the McGovern Institute for Brain Research at MIT, and the Center for Brains, Minds and Machines (CBMM), funded by an NSF STC award (CCF-1231216).

Author contributions: HLK and RS designed the study. HLK, LH, and IN collected the data. HLK analyzed the data with input from and supervision by MAC, NK, and RS. HLK, MAC, NK, and RS wrote the manuscript. All authors provided feedback on the final version.

612 **References**

613 [1] Jayaraman S, Fausey CM, Smith LB. Why are faces denser in the visual experiences of
614 younger than older infants? Dev Psychol 2017;53:38–49.
615 https://doi.org/10.1037/dev0000230.
616 [2] Jayaraman S, Fausey CM, Smith LB. The Faces in Infant-Perspective Scenes Change over the
617 First Year of Life. PLoS One 2015;10:e0123780.
618 https://doi.org/10.1371/journal.pone.0123780.
619 [3] Fausey CM, Jayaraman S, Smith LB. From faces to hands: Changing visual input in the first
620 two years. Cognition 2016;152:101–7. https://doi.org/10.1016/j.cognition.2016.03.005.
621 [4] Goren CC, Sarty M, Wu PYK. Visual following and pattern discrimination of face like stimuli
622 by newborn infants. Pediatrics 1975.
623 [5] Johnson MH, Dziurawiec S, Ellis H, Morton J. Newborns' preferential tracking of face-like
624 stimuli and its subsequent decline. Cognition 1991;40:1–19. https://doi.org/10.1016/0010-
625 0277(91)90045-6.
626 [6] Farroni T, Johnson MH, Menon E, Zulian L, Faraguna D, Csibra G. Newborns' preference
627 for face-relevant stimuli: Effects of contrast polarity. Proc Natl Acad Sci 2005;102:17245–50.
628 https://doi.org/10.1073/pnas.0502205102.
629 [7] Mondloch CJ, Lewis TL, Budreau DR, Maurer D, Dannemiller JL, Stephens BR, et al. Face
630 perception during early infancy. Psychol Sci 1999;10:419–22. https://doi.org/10.1111/1467-
631 9280.00179.
632 [8] Valenza E, Simion F, Cassia VM, Umiltà C. Face Preference at Birth. J Exp Psychol Hum
633 Percept Perform 1996;22:892–903. https://doi.org/10.1037/0096-1523.22.4.892.
634 [9] Turati C, Simion F, Milani I, Umiltà C. Newborns' preference for faces: what is crucial? Dev
635 Psychol 2002;38:875–82. https://doi.org/10.1037/0012-1649.38.6.875.
636 [10] Bushnell IWR. Mother's Face Recognition in Newborn Infants: Learning and Memory.
637 Infant Child Dev 2001;10:67–74. https://doi.org/10.1002/icd.248.
638 [11] Pascalis O, de Schonen S, Morton J, Deruelle C, Fabre-Grenet M. Mother's face recognition
639 by neonates: A replication and an extension. Infant Behav Dev 1995;18:79–85.
640 https://doi.org/10.1016/0163-6383(95)90009-8.
641 [12] Kosakowski HL, Cohen MA, Takahashi A, Keil B, Kanwisher N, Saxe R. Selective responses
642 to faces, scenes, and bodies in the ventral visual pathway of infants. Curr Biol 2021:1–10.
643 https://doi.org/10.1016/j.cub.2021.10.064.

[13] Kanwisher N, McDermott J, Chun MM. The Fusiform Face Area: A Module in Human Extrastriate Cortex Specialized for Face Perception. J Neurosci 1997;17:4302–11. https://doi.org/10.1523/JNEUROSCI.17-11-04302.1997.

[14] Gauthier I, Tarr MJ, Moylan J, Skudlarski P, Gore JC, Anderson AW. The fusiform "face area" is part of a network that processes faces at the individual level. J Cogn Neurosci 2000. https://doi.org/10.1162/089892900562165.

[15] Kanwisher N. Functional specificity in the human brain: A window into the functional architecture of the mind. Proc Natl Acad Sci 2010;107:11163–70. https://doi.org/10.1073/pnas.1005062107.

[16] Skerry AE, Saxe R. A common neural code for perceived and inferred emotion. J Neurosci 2014;34:15997–6008. https://doi.org/10.1523/JNEUROSCI.1676-14.2014.

[17] Deen B, Koldewyn K, Kanwisher N, Saxe R. Functional organization of social perception and cognition in the superior temporal sulcus. Cereb Cortex 2015;25:4596–609. https://doi.org/10.1093/cercor/bhv111.

[18] Lieberman MD, Straccia MA, Meyer ML, Du M, Tan KM. Social, self, (situational), and affective processes in medial prefrontal cortex (MPFC): Causal, multivariate, and reverse inference evidence. Neurosci Biobehav Rev 2019;99:311–28. https://doi.org/10.1016/j.neubiorev.2018.12.021.

[19] Powell LJ, Kosakowski HL, Saxe R. Social Origins of Cortical Face Areas. Trends Cogn Sci 2018;22:752–63. https://doi.org/10.1016/j.tics.2018.06.009.

[20] Arcaro MJ, Livingstone MS. On the relationship between maps and domains in inferotemporal cortex. Nat Rev Neurosci 2021. https://doi.org/10.1038/s41583-021-00490-4.

[21] Ciarrusta J, Dimitrova R, McAlonan G. Early maturation of the social brain: How brain development provides a platform for the acquisition of social-cognitive competence. vol. 254. 1st ed. Elsevier B.V.; 2020. https://doi.org/10.1016/bs.pbr.2020.05.004.

[22] Morton J, Johnson MH. CONSPEC and CONLERN: A two-process theory of infant face recognition. Psychol Rev 1991;98:164–81. https://doi.org/10.1037/0033-295X.98.2.164.

[23] Kanwisher N. Neural events and perceptual awareness. Cognition 2001;79:89–113. https://doi.org/10.1016/S0010-0277(00)00125-6.

[24] Arcaro MJ, Schade PF, Livingstone MS. Universal Mechanisms and the Development of the Face Network: What You See Is What You Get. Annu Rev Vis Sci 2019;5.

676 https://doi.org/10.1146/annurev-vision-091718-014917.
677 [25] Yue X, Robert S, Ungerleider LG. Curvature processing in human visual cortical areas.
678 Neuroimage 2020;222. https://doi.org/10.1016/j.neuroimage.2020.117295.
679 [26] Yue X, Pourladian IS, Tootell RBH, Ungerleider LG. Curvature-processing network in
680 macaque visual cortex. Proc Natl Acad Sci U S A 2014;111.
681 https://doi.org/10.1073/pnas.1412616111.
682 [27] Arcaro MJ, Schade PF, Vincent JL, Ponce CR, Livingstone MS. Seeing faces is necessary for
683 face-domain formation. Nat Neurosci 2017;20:1404–12. https://doi.org/10.1038/nn.4635.
684 [28] Puce A, Allison T, Bentin S, Gore JC, McCarthy G. Temporal cortex activation in humans
685 viewing eye and mouth movements. J Neurosci 1998;18:2188–99.
686 https://doi.org/10.1523/jneurosci.18-06-02188.1998.
687 [29] Isik L, Koldewyn K, Beeler D, Kanwisher N. Perceiving social interactions in the posterior
688 superior temporal sulcus. Proc Natl Acad Sci 2017;114:E9145–52.
689 https://doi.org/10.1073/pnas.1714471114.
690 [30] Pernet CR, McAleer P, Latinus M, Gorgolewski KJ, Charest I, Bestelmeyer PEG, et al. The
691 human voice areas: Spatial organization and inter-individual variability in temporal and extra-
692 temporal cortices. Neuroimage 2015;119:164–74.
693 https://doi.org/10.1016/j.neuroimage.2015.06.050.
694 [31] Belin P, Zatorre RJ, Lafaille P, Ahad P, Pike B. Voice-Selective Areas in Human Auditory
695 Cortex. Found Soc Neurosci 2018:309–12.
696 https://doi.org/10.7551/mitpress/3077.003.0023.
697 [32] Deen B, Saxe R, Kanwisher N. Processing communicative facial and vocal cues in the
698 superior temporal sulcus. Neuroimage 2020;221.
699 https://doi.org/10.1016/j.neuroimage.2020.117191.
700 [33] Beauchamp MS, Argall BD, Bodurka J, Duyn JH, Martin A. Unraveling multisensory
701 integration: Patchy organization within human STS multisensory cortex. Nat Neurosci
702 2004;7:1190–2. https://doi.org/10.1038/nn1333.
703 [34] Harris LT, McClure SM, Van Den Bos W, Cohen JD, Fiske ST. Regions of the MPFC
704 differentially tuned to social and nonsocial affective evaluation. Cogn Affect Behav Neurosci
705 2007;7:309–16. https://doi.org/10.3758/CABN.7.4.309.
706 [35] DiNicola LM, Braga RM, Buckner RL. Parallel distributed networks dissociate episodic and
707 social functions within the individual. J Neurophysiol 2020;123:1144–79.

708 https://doi.org/10.1152/JN.00529.2019.
709 [36] Shamay-Tsoory SG, Tomer R, Berger BD, Goldsher D, Aharon-Peretz J. Impaired "affective
710 theory of mind" is associated with right ventromedial prefrontal damage. Cogn Behav Neurol
711 2005;18:55–67. https://doi.org/10.1097/01.wnn.0000152228.90129.99.
712 [37] Moran JM, Macrae CN, Heatherton TF, Wyland CL, Kelley WM. Neuroanatomical Evidence
713 for Distinct Cognitive and Affective Components of Self. J Cogn Neurosci 2006;18:1586–94.
714 [38] D'Argembeau A, Feyers D, Majerus S, Collette F, Van der Linden M, Maquet P, et al. Self-
715 reflection across time: Cortical midline structures differentiate between present and past
716 selves. Soc Cogn Affect Neurosci 2008;3:244–52. https://doi.org/10.1093/scan/nsn020.
717 [39] D'Argembeau A, Collette F, Van Der Linden M, Laureys S, Del Fiore G, Degueldre C, et al.
718 Self-referential reflective activity and its relationship with rest: A PET study. Neuroimage
719 2005;25:616–24. https://doi.org/10.1016/j.neuroimage.2004.11.048.
720 [40] Farb NAS, Segal Z V., Mayberg H, Bean J, Mckeon D, Fatima Z, et al. Attending to the
721 present: Mindfulness meditation reveals distinct neural modes of self-reference. Soc Cogn
722 Affect Neurosci 2007;2:313–22. https://doi.org/10.1093/scan/nsm030.
723 [41] Jenkins AC, Macrae CN, Mitchell JP. Repetition suppression of ventromedial prefrontal
724 activity during judgments of self and others. Proc Natl Acad Sci U S A 2008;105:4507–12.
725 https://doi.org/10.1073/pnas.0708785105.
726 [42] Peelen M V., Atkinson AP, Vuilleumier P. Supramodal representations of perceived emotions
727 in the human brain. J Neurosci 2010;30:10127–34.
728 https://doi.org/10.1523/JNEUROSCI.2161-10.2010.
729 [43] Bruneau E, Dufour N, Saxe R. How We Know It Hurts: Item Analysis of Written Narratives
730 Reveals Distinct Neural Responses to Others' Physical Pain and Emotional Suffering. PLoS
731 One 2013;8. https://doi.org/10.1371/journal.pone.0063085.
732 [44] Bruneau EG, Pluta A, Saxe R. Distinct roles of the "Shared Pain" and "Theory of Mind"
733 networks in processing others' emotional suffering. Neuropsychologia 2012;50:219–31.
734 https://doi.org/10.1016/j.neuropsychologia.2011.11.008.
735 [45] Jacoby N, Bruneau E, Koster-Hale J, Saxe R. Localizing Pain Matrix and Theory of Mind
736 networks with both verbal and non-verbal stimuli. Neuroimage 2016.
737 https://doi.org/10.1016/j.neuroimage.2015.11.025.
738 [46] Shinkareva S V., Gao C, Wedell D. Audiovisual Representations of Valence: a Cross-study
739 Perspective. Affect Sci 2020;1:237–46. https://doi.org/10.1007/s42761-020-00023-9.

740 [47] Skerry AE, Saxe R. Neural Representations of Emotion Are Organized around Abstract
741 Event Features. Curr Biol 2015;25:1945–54. https://doi.org/10.1016/j.cub.2015.06.009.
742 [48] Brody BA, Kinney HC, Kloman AS, Gilles FH. Sequence of central nervous system
743 myelination in human infancy. J Neuropathol Exp Neurol 1987.
744 [49] Tau GZ, Peterson BS. Normal development of brain circuits. Neuropsychopharmacology
745 2010;35:147–68. https://doi.org/10.1038/npp.2009.115.
746 [50] Vasung L, Abaci Turk E, Ferradal SL, Sutin J, Stout JN, Ahtam B, et al. Exploring early
747 human brain development with structural and physiological neuroimaging. Neuroimage
748 2019;187:226–54. https://doi.org/10.1016/j.neuroimage.2018.07.041.
749 [51] Smyser CD, Snyder AZ, Neil JJ. Functional connectivity MRI in infants: Exploration of the
750 functional organization of the developing brain. Neuroimage 2011;56:1437–52.
751 https://doi.org/10.1016/j.neuroimage.2011.02.073.
752 [52] Dubois J, Poupon C, Thirion B, Simonnet H, Kulikova S, Leroy F, et al. Exploring the Early
753 Organization and Maturation of Linguistic Pathways in the Human Infant Brain. Cereb
754 Cortex 2016;26:2283–98. https://doi.org/10.1093/cercor/bhv082.
755 [53] Johnson MH. Cortical Maturation and the Development of Visual Attention in Early Infancy.
756 J Cogn Neurosci 1990;2:81–95. https://doi.org/10.1162/jocn.1990.2.2.81.
757 [54] Sydnor VJ, Larsen B, Bassett DS, Alexander-bloch A, Fair DA, Liston C, et al.
758 Neurodevelopment of the association cortices: Patterns, mechanisms, and implications for
759 psychopathology. Neuron 2021:1–27. https://doi.org/10.1016/j.neuron.2021.06.016.
760 [55] Li G, Nie J, Wang L, Shi F, Lin W, Gilmore JH, et al. Mapping region-specific longitudinal
761 cortical surface expansion from birth to 2 years of age. Cereb Cortex 2013;23:2724–33.
762 https://doi.org/10.1093/cercor/bhs265.
763 [56] Meng Y, Li G, Lin W, Gilmore JH, Shen D. Spatial distribution and longitudinal
764 development of deep cortical sulcal landmarks in infants. Neuroimage 2014;100:206–18.
765 https://doi.org/10.1016/j.neuroimage.2014.06.004.
766 [57] Miller DJ, Duka T, Stimpson CD, Schapiro SJ, Baze WB, McArthur MJ, et al. Prolonged
767 myelination in human neocortical evolution. Proc Natl Acad Sci U S A 2012;109:16480–5.
768 https://doi.org/10.1073/pnas.1117943109.
769 [58] Carmody DP, Dunn SM, Boddie-Willis AS, DeMarco JK, Lewis M. A quantitative measure of
770 myelination development in infants, using MR images. Neuroradiology 2004;46:781–6.
771 https://doi.org/10.1007/s00234-004-1241-z.

[59] Hasegawa M, Houdou S, Mito T, Takashima S, Asanuma K, Ohno T. Development of myelination in the human fetal and infant cerebrum: A myelin basic protein immunohistochemical study. Brain Dev 1992;14:1–6. https://doi.org/10.1016/S0387-7604(12)80271-3.

[60] Chugani HT, Phelps ME. Maturational changes in cerebral function in infants determined by 18FDG positron emission tomography. Science (80-) 1986;231:840–3. https://doi.org/10.1126/science.3945811.

[61] Huttenlocher PR, Dabholkar AS. Regional differences in synaptogenesis in human cerebral cortex. J Comp Neurol 1997;387:167–78. https://doi.org/10.1002/(SICI)1096-9861(19971020)387:2<167::AID-CNE1>3.0.CO;2-Z.

[62] Sanai N, Nguyen T, Ihrie RA, Mirzadeh Z, Tsai HH, Wong M, et al. Corridors of migrating neurons in the human brain and their decline during infancy. Nature 2011;478:382–6. https://doi.org/10.1038/nature10487.

[63] Kostović I, Sedmak G, Judaš M. Neural histology and neurogenesis of the human fetal and infant brain. Neuroimage 2019;188:743–73. https://doi.org/10.1016/j.neuroimage.2018.12.043.

[64] Bushnell IWR, Sai F, Mullin JT. Neonatal recognition of the mother's face. Br J Dev Psychol 1989;7:3–15. https://doi.org/10.1111/j.2044-835X.1989.tb00784.x.

[65] Field TM, Cohen D, Garcia R, Greenberg R. Mother-stranger face discrimination by the newborn. Infant Behav Dev 1984;7:19–25. https://doi.org/10.1016/S0163-6383(84)80019-3.

[66] Kelly DJ, Quinn PC, Slater AM, Lee K, Gibson A, Smith M, et al. Three-month-olds, but not newborns, prefer own-race faces. Dev Sci 2005;8:31–6. https://doi.org/10.1111/j.1467-7687.2005.0434a.x.Three-month-olds.

[67] Ramsey-Rennels JL, Langlois JH. Infants' differential processing of female and male faces. Curr Dir Psychol Sci 2006;15:59–62. https://doi.org/10.1111/j.0963-7214.2006.00407.x.

[68] Turati C, Montirosso R, Brenna V, Ferrara V, Borgatti R. A Smile Enhances 3-Month-Olds' Recognition of an Individual Face. Infancy 2011;16:306–17. https://doi.org/10.1111/j.1532-7078.2010.00047.x.

[69] Bayet L, Quinn PC, Tanaka JW, Lee K, Gentaz É, Pascalis O. Face gender influences the looking preference for smiling expressions in 3.5-month-old human infants. PLoS One 2015;10:6–13. https://doi.org/10.1371/journal.pone.0129812.

[70] Kiley Hamlin J, Wynn K, Bloom P. Three-month-olds show a negativity bias in their social

804 evaluations. Dev Sci 2010;13:923–9. https://doi.org/10.1111/j.1467-7687.2010.00951.x.
805 [71] Kinzler KD, Dupoux E, Spelke ES. The native language of social cognition. Proc Natl Acad
806 Sci 2007;104:12577–80. https://doi.org/10.1073/pnas.0705345104.
807 [72] Powell LJ, Spelke ES. Human infants' understanding of social imitation: Inferences of
808 affiliation from third party observations. Cognition 2018;170:31–48.
809 https://doi.org/10.1016/j.cognition.2017.09.007.
810 [73] Johnson SC, Dweck CS, Chen FS, Stern HL, Ok SJ, Barth M. At the intersection of social
811 and cognitive development: Internal working models of attachment in infancy. Cogn Sci
812 2010;34:807–25. https://doi.org/10.1111/j.1551-6709.2010.01112.x.
813 [74] Farroni T, Chiarelli AM, Lloyd-Fox S, Massaccesi S, Merla A, Di Gangi V, et al. Infant cortex
814 responds to other humans from shortly after birth. Sci Rep 2013;3:3–7.
815 https://doi.org/10.1038/srep02851.
816 [75] Ichikawa H, Kanazawa S, Yamaguchi MK, Kakigi R. Infant brain activity while viewing facial
817 movement of point-light displays as measured by near-infrared spectroscopy (NIRS).
818 Neurosci Lett 2010;482:90–4. https://doi.org/10.1016/j.neulet.2010.06.086.
819 [76] Powell LJ, Deen B, Saxe R. Using individual functional channels of interest to study cortical
820 development with fNIRS. Dev Sci 2018;21:e12595. https://doi.org/10.1111/desc.12595.
821 [77] Deen B, Richardson H, Dilks DD, Takahashi A, Keil B, Wald LL, et al. Organization of high-
822 level visual cortex in human infants. Nat Commun 2017;8:1–10.
823 https://doi.org/10.1038/ncomms13995.
824 [78] de Heering A, Rossion B. Rapid categorization of natural face images in the infant right
825 hemisphere. Elife 2015;4:1–14. https://doi.org/10.7554/eLife.06564.
826 [79] Lloyd-Fox S, Begus K, Halliday D, Pirazzoli L, Blasi A, Papademetriou M, et al. Cortical
827 specialisation to social stimuli from the first days to the second year of life: A rural Gambian
828 cohort. Dev Cogn Neurosci 2017;25:92–104. https://doi.org/10.1016/j.dcn.2016.11.005.
829 [80] Tzourio-Mazoyer N, De Schonen S, Crivello F, Reutter B, Aujard Y, Mazoyer B. Neural
830 correlates of woman face processing by 2-month-old infants. Neuroimage 2002;15:454–61.
831 https://doi.org/10.1006/nimg.2001.0979.
832 [81] Grossmann T, Johnson MH, Lloyd-Fox S, Blasi A, Deligianni F, Elwell C, et al. Early cortical
833 specialization for face-to-face communication in human infants. Proc R Soc B Biol Sci
834 2008;275:2803–11. https://doi.org/10.1098/rspb.2008.0986.
835 [82] Krol KM, Grossmann T. Impression Formation in the Human Infant Brain. Cereb Cortex

836 Commun 2020:1–11. https://doi.org/10.1093/texcom/tgaa070.
837 [83] Keil B, Alagappan V, Mareyam A, McNab JA, Fujimoto K, Tountcheva V, et al. Size-
838 optimized 32-channel brain arrays for 3 T pediatric imaging. Magn Reson Med 2011;66:1777–
839 87. https://doi.org/10.1002/mrm.22961.
840 [84] Zapp J, Schmitter S, Schad LR. Sinusoidal echo-planar imaging with parallel acquisition
841 technique for reduced acoustic noise in auditory fMRI. J Magn Reson Imaging 2012;36:581–
842 8. https://doi.org/10.1002/jmri.23699.
843 [85] Ghotra A, Kosakowski HL, Takahashi A, Etzel R, May MW, Scholz A, et al. A size-adaptive
844 32-channel array coil for awake infant neuroimaging at 3 Tesla MRI. Magn Reson Med
845 2021;86:1773–85. https://doi.org/10.1002/mrm.28791.
846 [86] Cousineau D. Confidence intervals in within-subject designs: A simpler solution to Loftus
847 and Masson's method. Tutor Quant Methods Psychol 2005;1:42–5.
848 [87] Pitcher D, Dilks DD, Saxe RR, Triantafyllou C, Kanwisher N. Differential selectivity for
849 dynamic versus static information in face-selective cortical regions. Neuroimage
850 2011;56:2356–63. https://doi.org/10.1016/j.neuroimage.2011.03.067.
851 [88] Adamson LB, Frick JE. The still face: A history of a shared experimental paradigm. Infancy
852 2003;4:451–73. https://doi.org/10.1207/S15327078IN0404_01.
853 [89] Toda S, Fogel A. Infant response to the still-face situation at 3 and 6 months. Dev Psychol
854 1993;29:532–8. https://doi.org/10.1037/0012-1649.29.3.532.
855 [90] Lloyd-Fox S, Blasi A, Everdell N, Elwell CE, Johnson MH. Selective cortical mapping of
856 biological motion processing in young infants. J Cogn Neurosci 2011;23:2521–32.
857 https://doi.org/10.1162/jocn.2010.21598.
858 [91] Lloyd-Fox S, Blasi A, Volein A, Everdell N, Elwell CE, Johnson MH. Social perception in
859 infancy: A near infrared spectroscopy study. Child Dev 2009;80:986–99.
860 https://doi.org/10.1111/j.1467-8624.2009.01312.x.
861 [92] Poulin-Dubois D, Crivello C, Wright K. Biological motion primes the animate/inanimate
862 distinction in infancy. PLoS One 2015;10:1–12.
863 https://doi.org/10.1371/journal.pone.0116910.
864 [93] Powell LJ, Spelke ES. Human infants' understanding of social imitation: Inferences of
865 affiliation from third party observations. Cognition 2018;170:31–48.
866 https://doi.org/10.1016/j.cognition.2017.09.007.
867 [94] Hamlin JK, Wynn K, Bloom P. Social evaluation by preverbal infants. Nature 2007;450:557–

9. https://doi.org/10.1038/nature06288.

[95] Liu S, Spelke ES. Six-month-old infants expect agents to minimize the cost of their actions. Cognition 2017;160:35–42. https://doi.org/10.1016/j.cognition.2016.12.007.

[96] Baker RK, Pettigrew TL, Poulin-Dubois D. Infants' ability to associate motion paths with object kinds. Infant Behav Dev 2014;37:119–29. https://doi.org/10.1016/j.infbeh.2013.12.005.

[97] Lisboa IC, Queirós S, Miguel H, Sampaio A, Santos JA, Pereira AF. Infants' cortical processing of biological motion configuration – A fNIRS study. Infant Behav Dev 2020. https://doi.org/10.1016/j.infbeh.2020.101450.

[98] Farroni T, Menon E, Rigato S, Johnson MH. The perception of facial expressions in newborns. Eur J Dev Psychol 2007;4:2–13. https://doi.org/10.1080/17405620601046832.

[99] Addabbo M, Longhi E, Marchis IC, Tagliabue P, Turati C. Dynamic facial expressions of emotions are discriminated at birth. PLoS One 2018;13:1–11. https://doi.org/10.1371/journal.pone.0193868.

[100] Quinn PC, Uttley L, Lee K, Gibson A, Smith M, Slater AM, et al. Infant preference for female faces occurs for same- but not other-race faces. J Neuropsychol 2008;2:15–26. https://doi.org/10.1348/174866407X231029.

[101] Quinn PC, Yahr J, Kuhn A, Slater AM, Pascalis O. Representation of the Gender of Human Faces by Infants: A Preference for Female. Perception 2002;31:1109–21. https://doi.org/10.1068/p3331.

[102] Ramsey JL, Langlois JH, Marti NC. Infant categorization of faces: Ladies first. Dev Rev 2005;25:212–46. https://doi.org/10.1016/j.dr.2005.01.001.

[103] Tham DSY, Bremmer JG, Hay D. In Infancy the Timing of Emergence of the Other-Race-Effect is Dependent on Face Gender. Infant Behav Dev 2015;40:131–8.

[104] Liu S, Xiao NG, Quinn PC, Zhu D, Ge L, Pascalis O, et al. Asian infants show preference for own-race but not other-race female faces: The role of infant caregiving arrangements. Front Psychol 2015;6:1–8. https://doi.org/10.3389/fpsyg.2015.00593.

[105] Mehr SA, Song LA, Spelke ES. For 5-Month-Old Infants, Melodies Are Social. Psychol Sci 2016;27:486–501. https://doi.org/10.1177/0956797615626691.

[106] Bayet L, Perdue KL, Behrendt HF, Richards JE, Westerlund A, Cataldo JK, et al. Neural responses to happy, fearful and angry faces of varying identities in 5- and 7-month-old infants. Dev Cogn Neurosci 2021;47:100882. https://doi.org/10.1016/j.dcn.2020.100882.

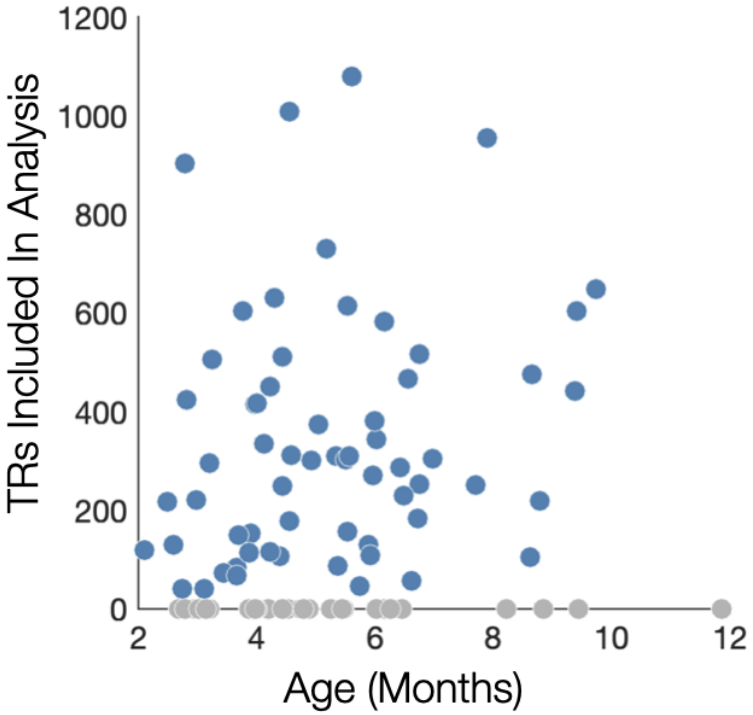
900 [107] Nakato E, Otsuka Y, Kanazawa S, Yamaguchi MK, Kakigi R. Distinct differences in the
901 pattern of hemodynamic response to happy and angry facial expressions in infants - A near-
902 infrared spectroscopic study. Neuroimage 2011;54:1600–6.
903 https://doi.org/10.1016/j.neuroimage.2010.09.021.
904 [108] Saito Y, Aoyama S, Kondo T, Fukumoto R, Konishi N, Nakamura K, et al. Frontal cerebral
905 blood flow change associated with infant-directed speech. Arch Dis Child Fetal Neonatal Ed
906 2007;92. https://doi.org/10.1136/adc.2006.097949.
907 [109] Kotilahti K, Nissilä I, Näsi T, Lipiäinen L, Noponen T, Meriläinen P, et al. Hemodynamic
908 responses to speech and music in newborn infants. Hum Brain Mapp 2010;31:595–603.
909 https://doi.org/10.1002/hbm.20890.
910 [110] Sato H, Hirabayashi Y, Tsubokura H, Kanai M, Ashida T, Konishi I, et al. Cerebral
911 hemodynamics in newborn infants exposed to speech sounds: A whole-head optical
912 topography study. Hum Brain Mapp 2012;33:2092–103.
913 https://doi.org/10.1002/hbm.21350.
914 [111] Jönsson EH, Kotilahti K, Heiskala J, Wasling HB, Olausson H, Croy I, et al. Affective and
915 non-affective touch evoke differential brain responses in 2-month-old infants. Neuroimage
916 2018;169:162–71. https://doi.org/10.1016/j.neuroimage.2017.12.024.
917 [112] Báez-Mendoza R, Mastrobattista EP, Wang AJ, Williams ZM. Social agent identity cells in the
918 prefrontal cortex of interacting groups of primates. Science (80-) 2021;374.
919 https://doi.org/10.1126/science.abb4149.
920 [113] Haber SN, Knutson B. The reward circuit: Linking primate anatomy and human imaging.
921 Neuropsychopharmacology 2010;35:4–26. https://doi.org/10.1038/npp.2009.129.
922 [114] Pastor V, Medina JH. Medial prefrontal cortical control of reward- and aversion-based
923 behavioral output: Bottom-up modulation. Eur J Neurosci 2021;53:3039–62.
924 https://doi.org/10.1111/ejn.15168.
925 [115] Bhanji JP, Delgado MR. The social brain and reward: Social information processing in the
926 human striatum. Wiley Interdiscip Rev Cogn Sci 2014;5:61–73.
927 https://doi.org/10.1002/wcs.1266.
928 [116] Della Longa L, Gliga T, Farroni T. Tune to touch: Affective touch enhances learning of face
929 identity in 4-month-old infants. Dev Cogn Neurosci 2019;35:42–6.
930 https://doi.org/10.1016/j.dcn.2017.11.002.
931 [117] Della Longa L, Carnevali L, Patron E, Dragovic D, Farroni T. Psychophysiological and

932 Visual Behavioral Responses to Faces Associated with Affective and Non-affective Touch in
933 Four-month-old Infants. Neuroscience 2021;464:67–78.
934 https://doi.org/10.1016/j.neuroscience.2020.07.053.
935 [118] Naoi N, Minagawa-Kawai Y, Kobayashi A, Takeuchi K, Nakamura K, Yamamoto J ichi, et al.
936 Cerebral responses to infant-directed speech and the effect of talker familiarity. Neuroimage
937 2012;59:1735–44. https://doi.org/10.1016/j.neuroimage.2011.07.093.
938 [119] Sowell ER, Peterson BS, Thompson PM, Welcome SE, Henkenius AL, Toga AW. Mapping
939 cortical change across the human life span. Nat Neurosci 2003;6:309–15.
940 https://doi.org/10.1038/nn1008.
941 [120] Casey BJ, Jones RM, Hare TA. The Adolescent Brain. Ann N Y Acad Sci 1008;1124:111–26.
942 [121] Casey BJ, Trainor RJ, Orendi JL, Schubert AB, Nystrom LE, Giedd JN, et al. A
943 developmental functional MRI study of prefrontal activation during performance of a Go-
944 No-Go task. J Cogn Neurosci 1997;9:835–47. https://doi.org/10.1162/jocn.1997.9.6.835.
945 [122] Grossmann T. The role of medial prefrontal cortex in early social cognition. Front Hum
946 Neurosci 2013;7. https://doi.org/10.3389/fnhum.2013.00340.
947 [123] Werchan DM, Collins AGE, Frank MJ, Amso D. Role of prefrontal cortex in learning and
948 generalizing hierarchical rules in 8-month-old infants. J Neurosci 2016;36:10314–22.
949 https://doi.org/10.1523/JNEUROSCI.1351-16.2016.
950 [124] Ellis CT, Skalaban LJ, Yates TS, Turk-Browne NB. Attention recruits frontal cortex in
951 human infants. Proc Natl Acad Sci U S A 2021;118.
952 https://doi.org/10.1073/pnas.2021474118.
953 [125] Gervain J, Macagno F, Cogoi S, Peña M, Mehler J. The neonate brain detects speech
954 structure. Proc Natl Acad Sci U S A 2008;105:14222–7.
955 https://doi.org/10.1073/pnas.0806530105.
956 [126] Dehaene-Lambertz G, Montavont A, Jobert A, Allirol L, Dubois J, Hertz-Pannier L, et al.
957 Language or music, mother or Mozart? Structural and environmental influences on infants'
958 language networks. Brain Lang 2010;114:53–65.
959 https://doi.org/10.1016/j.bandl.2009.09.003.
960 [127] Vouloumanos A, Hauser MD, Werker JF, Martin A. The tuning of human neonates'
961 preference for speech. Child Dev 2010;81:517–27. https://doi.org/10.1111/j.1467-
962 8624.2009.01412.x.
963 [128] Minagawa-kawai Y, Lely H Van Der, Ramus F, Sato Y, Mazuka R, Dupoux E. Optical Brain

964 Imaging Reveals General Auditory and Language-Specific Processing in Early Infant
965 Development 2010. https://doi.org/10.1093/cercor/bhq082.
966 [129] Altvater-Mackensen N, Grossmann T. Modality-independent recruitment of inferior frontal
967 cortex during speech processing in human infants. Dev Cogn Neurosci 2018;34:130–8.
968 https://doi.org/10.1016/j.dcn.2018.10.002.
969 [130] May L, Byers-Heinlein K, Gervain J, Werker JF. Language and the newborn brain: Does
970 prenatal language experience shape the neonate neural response to speech? Front Psychol
971 2011;2:1–9. https://doi.org/10.3389/fpsyg.2011.00222.
972 [131] Li J, Osher DE, Hansen HA, Saygin ZM. Innate connectivity patterns drive the development
973 of the visual word form area. Sci Rep 2020;10:1–12. https://doi.org/10.1038/s41598-020-
974 75015-7.
975 [132] Werchan DM, Amso D. A Novel Ecological Account of Prefrontal Cortex Functional
976 Development. Physchol Rev 2017;124:720–39.
977 [133] Dehaene-Lambertz G, Spelke ES. The Infancy of the Human Brain. Neuron 2015;88:93–109.
978 https://doi.org/10.1016/j.neuron.2015.09.026.
979 [134] Saygin ZM, Osher DE, Norton ES, Youssoufian DA, Beach SD, Feather J, et al.
980 Connectivity precedes function in the development of the visual word form area. Nat
981 Neurosci 2016;19:1250–5. https://doi.org/10.1038/nn.4354.
982 [135] Dweck CS. From needs to goals and representations: Foundations for a unified theory of
983 motivation, personality, and development. Psychol Rev 2017;124:689–719.
984 https://doi.org/10.1037/rev0000082.
985 [136] Kay KN, Rokem A, Winawer J, Dougherty RF, Wandell BA. GLMdenoise: A fast, automated
986 technique for denoising task-based fMRI data. Front Neurosci 2013;7:1–15.
987 https://doi.org/10.3389/fnins.2013.00247.
988 [137] Arichi T, Fagiolo G, Varela M, Melendez-Calderon A, Allievi A, Merchant N, et al.
989 Development of BOLD signal hemodynamic responses in the human brain. Neuroimage
990 2012;63:663–73. https://doi.org/10.1016/j.neuroimage.2012.06.054.
991 [138] Glasser MF, Coalson TS, Robinson EC, Hacker CD, Harwell J, Yacoub E, et al. A multi-
992 modal parcellation of human cerebral cortex. Nature 2016;536:171–8.
993 https://doi.org/10.1038/nature18933.
994
995

996 # Supplemental Figures and Tables

997



998
999 **Figure S1. Quantity of low-motion TRs is not correlated with age.** The amount of data
1000 included for an individual data point was not correlated with the age of that subject (r=0.19,
1001 p=0.12). Each participant that contributed at least some low motion data is indicated by blue circles.
1002 Participants that contributed no data are indicated by gray points.
1003
1004

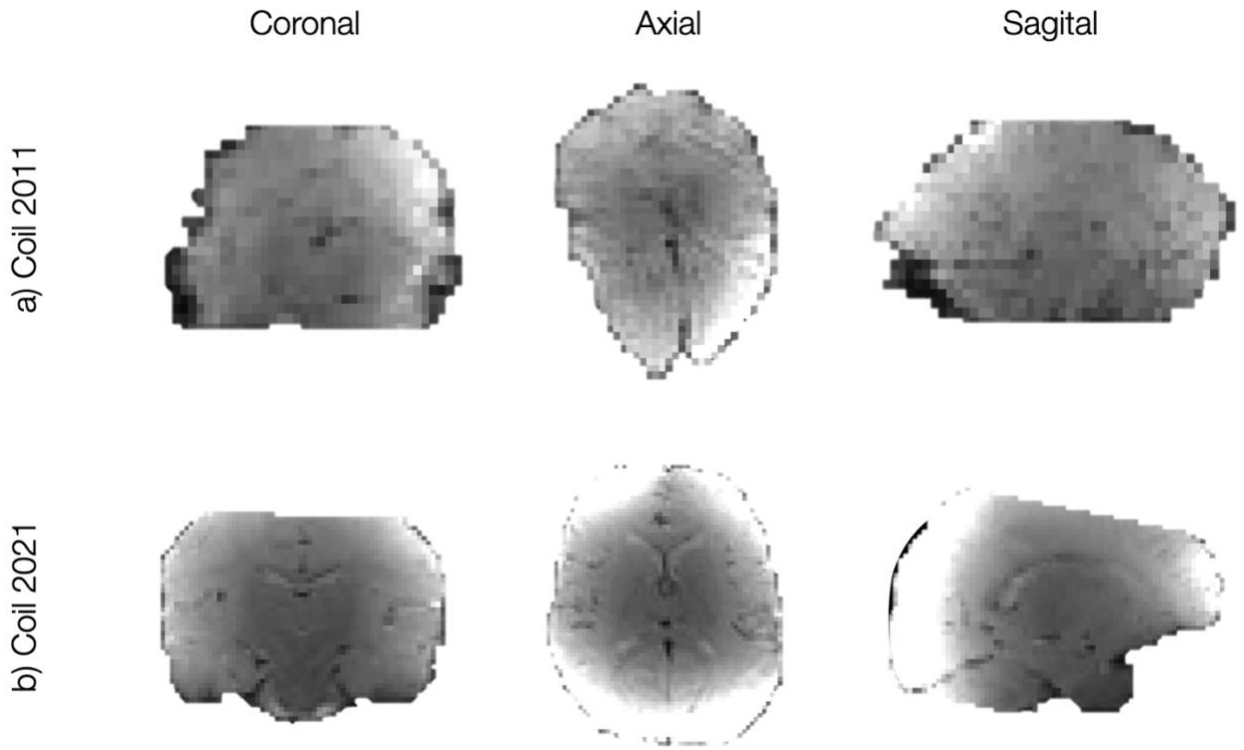


Figure S2. Distortions in infant fMRI data. Representative (not ideal) image of fMRI data collected on *Coil 2011* with EPI using a sinusoidal trajectory (a) and an EPI with a standard trajectory collected on *Coil 2021* (b).

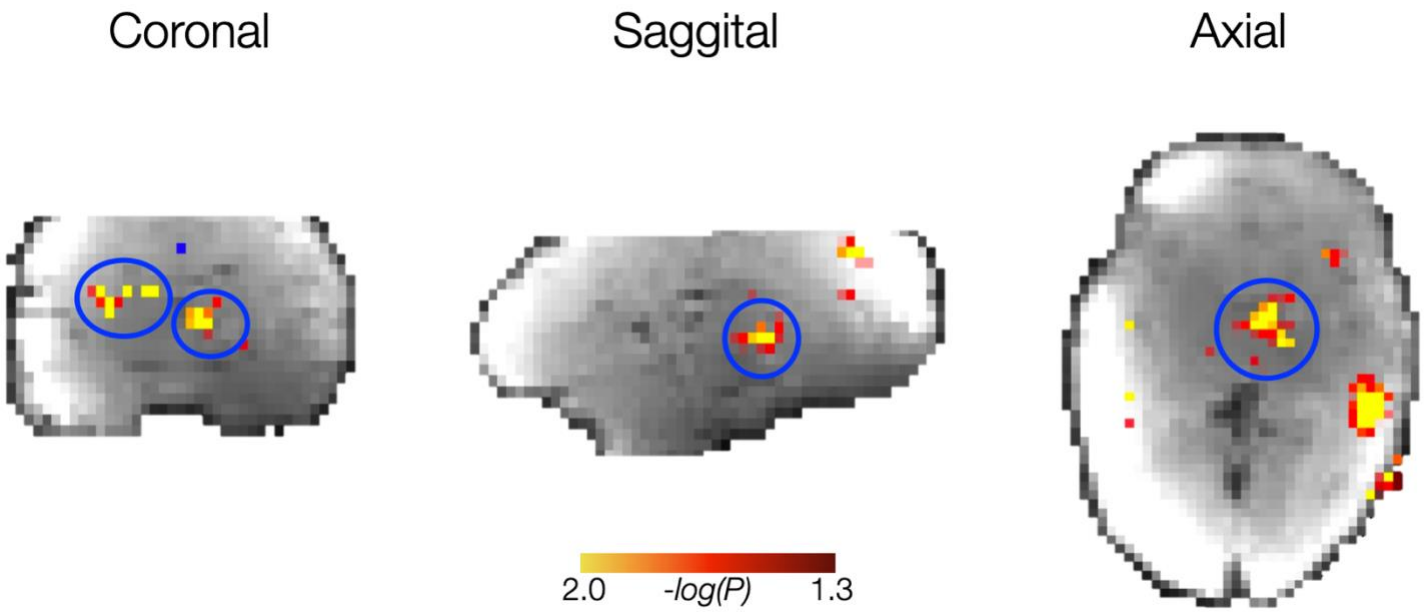


Figure S3. Subcortical responses to faces in *Coil 2011* dataset. A group random effects analysis of the *Coil 2011* dataset (n=26) revealed a subcortical activation to faces, outlined in blue. These face responses could not be localized to a specific subcortical area due to the distorted nature of the *Coil 2011* functional data.

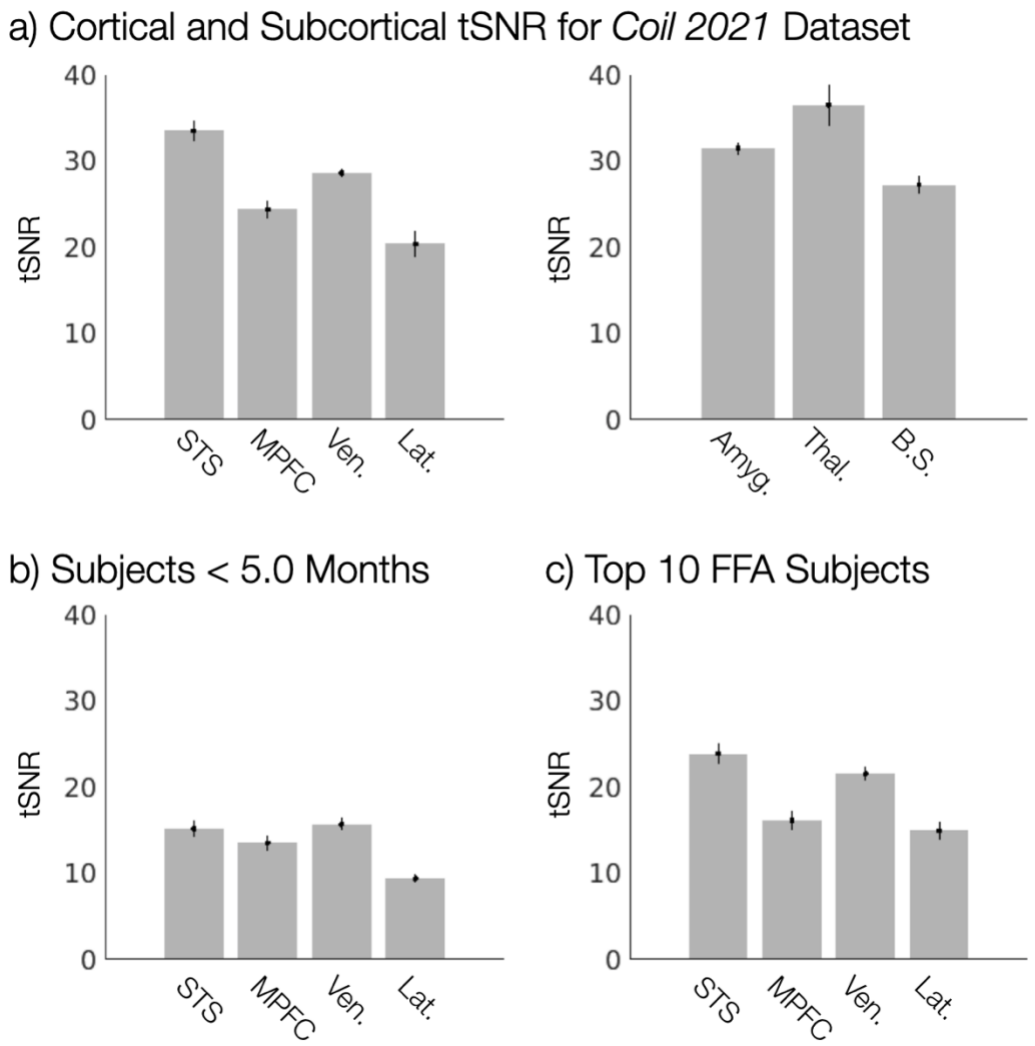


Figure S4. Temporal signal to noise (tSNR) in subsets of data. (a) tSNR across *Coil 2021* subjects in cortical (left) and subcortical (right) parcels. (b) tSNR for the youngest subjects across cortical parcels. (c) tSNR in cortical parcels for the 10 subjects with the most face selective responses. Error bars indicate within-subject SE.

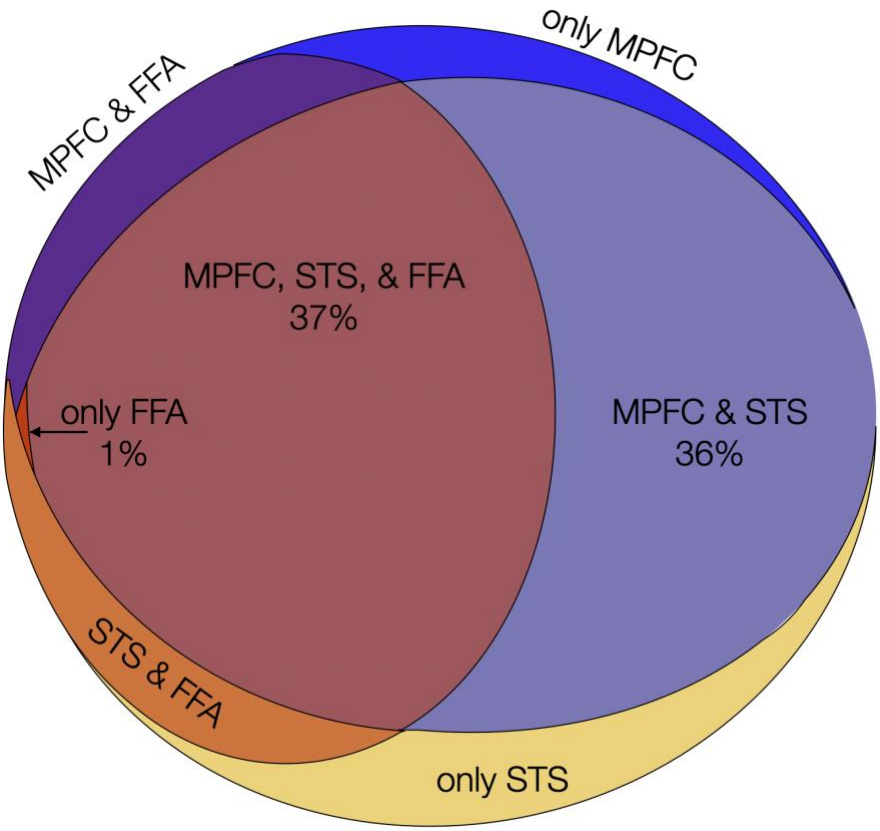


Figure S5. Permutation test reveals it is very unlikely that infants have face-selective responses in only FFA. Proportional Venn Diagram reveals the number of permutations out of 2500 that MPFC (blue), STS (yellow), and FFA (red) exhibited face-selective responses with overlapping observations represented by merged colors.

Table S1. Parameter estimates for face-selective responses.

Coil 2011	**Intercept¥**	**Age¥**	**Sex¥**	**Motion¥**	**Paradigm¥**	**F-B¥**	**F-O¥**	**F-S¥**
STS	**0.15**	-0.03	0.04	-0.01	0.06	**-0.16**	**-0.16**	**-0.15**
	(-0.03 - 0.34)	(-0.12 - 0.06)	(-0.20 - 0.29)	(-0.13 - 0.10)	(-0.17 - 0.29)	**(-0.31 - -0.02)**	**(-0.30 - -0.02)**	**(-0.30 - -0.01)**
MPFC	**0.44103**	-0.05	-0.20	**0.18**	0.13	**-0.31**	**-0.36**	**-0.36**
	(0.19 - 0.69)	(-0.17 - 0.07)	(-0.51 - 0.11)	**(0.04 - 0.33)**	(-0.16 - 0.42)	**(-0.53 - -0.09)**	**(-0.58 - -0.14)**	**(-0.58 - -0.14)**
Coil 2021								
STS	**0.46**	0.05	0.11	**0.17**	n/a	**-0.44**	**-0.44**	**-0.43**
	(0.27 - 0.65)	(-0.06 - 0.16)	(-0.11 - 0.33)	**(0.06 - 0.28)**		**(-0.63 - -0.25)**	**(-0.63 - -0.25)**	**(-0.62 - -0.25)**
MPFC	**0.40**	0.08	**0.34**	-0.09	n/a	**-0.60**	**-0.35**	**-0.50**
	(0.15 - 0.65)	(-0.07 - 0.23)	**(0.04 - 0.63)**	(-0.25 - 0.06)		**(-0.85 - -0.35)**	**(-0.60 - -0.10)**	**(-0.75 - -0.25)**
Amygdala	**0.16**	0.03	-0.11	0.05	n/a	**-0.20**	**-0.13**	-0.12
	(0.02 - 0.31)	(-0.06 - 0.11)	(-0.28 - 0.06)	(-0.03 - 0.14)		**(-0.35 - -0.05)**	**(-0.28 - 0.02)**	(-0.27 - 0.03)
Thalamus	**0.08**	0.01	-0.01	0.01	n/a	**-0.07**	**-0.08**	-0.04
	(0.01 - 0.15)	(-0.04 - 0.05)	(-0.09 - 0.07)	(-0.03 - 0.05)		**(-0.14 - -0.00)**	**(-0.15 - -0.01)**	(-0.11 - 0.03)
Brainstem	0.08	-0.02	-0.09	0.03	n/a	-0.05	-0.10	0.02
	(-0.02 - 0.18)	(-0.06 - 0.03)	(-0.18 - 0.00)	(-0.02 - 0.07)		(-0.18 - 0.08)	(-0.23 - 0.03)	(-0.10 - 0.15)
Combined								
STS	**0.37**	0.03	0.07	0.07	-0.09	**-0.28**	**-0.29**	**-0.29**
	(0.19 - 0.55)	(-0.05 - 0.10)	(-0.08 - 0.23)	(-0.01 - 0.15)	(-0.24 - 0.07)	**(-0.41 - -0.16)**	**(-0.41 - -0.17)**	**(-0.41 - -0.16)**
MPFC	**0.37**	0.03	0.20	0.01	-0.03	**-0.42**	**-0.34**	**-0.41**
	(0.10 - 0.64)	(-0.08 - 0.14)	(-0.03 - 0.42)	(-0.10 - 0.13)	(-0.25 - 0.20)	**(-0.60 - -0.25)**	**(-0.51 - -0.17)**	**(-0.59 - -0.24)**
Ventral	0.12	0.06	0.02	**-0.07**	0.10	**-0.29**	**-0.19**	**-0.28**
	(-0.05 - 0.28)	(-0.00 - 0.13)	(-0.11 - 0.15)	**(-0.14 - -0.01)**	(-0.04 - 0.24)	**(-0.43 - -0.15)**	**(-0.33 - -0.05)**	**(-0.41 - -0.14)**
Lateral	0.26	-0.09	-0.17	-0.13	0.21	**-0.30**	-0.17	**-0.42**
	(-0.08 - 0.59)	(-0.23 - 0.05)	(-0.44 - 0.11)	(-0.27 - 0.02)	(-0.07 - 0.49)	**(-0.51 - -0.08)**	(-0.38 - 0.05)	**(-0.64 - -0.21)**

¥ Parameter estimates from linear mixed-effects model with confidence interval in parenthesis; $P \leq 0.05$ indicated in bold; intercept is the difference between face videos and baseline videos; motion is the proportion of scrubbed voxels; paradigm indicates which paradigm was used.

1030
1031 **Table S2. Parameter Estimates for Linear Mixed Effects Model of Contrast Value.**

Cortical†	Intercept¥	Age¥	Sex¥	Motion¥
STS	**0.34**	0.00	-0.10	0.04
	(0.09 - 0.59)	(-0.13 - 0.13)	(-0.36 - 0.16)	(-0.09 - 0.18)
MPFC	**0.35**	0.04	0.09	-0.04
	(0.13 - 0.57)	(-0.12 - 0.20)	(-0.23 - 0.40)	(-0.20 - 0.12)
Ventral	**0.28**	0.05	-0.06	-0.02
	(0.07 - 0.49)	(-0.08 - 0.19)	(-0.32 - 0.21)	(-0.16 - 0.11)
Lateral	**0.31**	0.13	-0.04	-0.20
	(-0.00 - 0.63)	(-0.09 - 0.36)	(-0.49 - 0.41)	(-0.42 - 0.03)
Subcortical†				
Amygdala	0.20	0.12	-0.10	**0.16**
	(-0.01 - 0.40)	(-0.03 - 0.28)	(-0.40 - 0.20)	**(0.00 - 0.31)**
Thalamus	0.10	0.04	-0.07	0.00
	(-0.00 - 0.21)	(-0.03 - 0.12)	(-0.23 - 0.08)	(-0.08 - 0.08)
Brainstem	0.05	0.07	-0.01	0.01
	(-0.07 - 0.16)	(-0.01 - 0.16)	(-0.18 - 0.16)	(-0.08 - 0.09)

¥ Parameter estimates from linear mixed-effects model with confidence interval in parenthesis; $P \leq 0.05$ indicated in bold; intercept (int) is the difference between baseline and the contrast value; motion is the proportion of scrubbed voxels.
† Cortical parcels are combined across *Coil 2011* and *Coil 2021* datasets; Subcortical parcels are the *Coil 2011* dataset only.

1032
1033

Table S3. Results from Linear Mixed Effects Models for subsets of infants.

Youngest	Intercept[¥]	Age[¥]	Sex[¥]	Motion[¥]	B[¥]	O[¥]	S[¥]
STS[œ]	**0.46**	-0.00	-0.04	0.07	**-0.40**	**-0.34**	**-0.36**
	(0.19 - 0.72)	(-0.15 - 0.14)	(-0.33 - 0.25)	(-0.08 - 0.21)	**(-0.59 - -0.20)**	**(-0.54 - -0.15)**	**(-0.55 - -0.16)**
MPFC	0.18	-0.11	0.28	0.03	**-0.22**	**-0.24**	**-0.39**
	(-0.22 - 0.58)	(-0.33 - 0.11)	(-0.17 - 0.73)	(-0.19 - 0.25)	**(-0.47 - 0.03)**	**(-0.49 - 0.00)**	**(-0.64 - -0.14)**
Ventral	0.04	0.10	0.01	**-0.14**	**-0.22**	**-0.18**	-0.14
	(-0.24 - 0.31)	(-0.02 - 0.23)	(-0.23 - 0.25)	**(-0.25 - -0.02)**	**(-0.44 - -0.01)**	**(-0.40 - 0.04)**	(-0.35 - 0.08)
Lateral	0.41	**0.41**	**-0.43**	-0.18	-0.12	0.03	-0.20
	(-0.26 - 1.07)	**(0.19 - 0.62)**	**(-0.84 - -0.03)**	(-0.37 - 0.02)	(-0.49 - 0.24)	(-0.33 - 0.40)	(-0.57 - 0.16)
Best FFA[†]							
STS	**0.47**	0.11	-0.14	0.00	**-0.32**	**-0.31**	**-0.24**
	(0.15 - 0.78)	(-0.01 - 0.24)	(-0.40 - 0.12)	(-0.12 - 0.12)	**(-0.54 - -0.10)**	**(-0.53 - -0.09)**	**(-0.46 - -0.02)**
MPFC	**0.53**	0.01	0.27	**0.18**	**-0.62**	**-0.43**	**-0.61**
	(0.17 - 0.89)	(-0.16 - 0.17)	(-0.09 - 0.63)	**(0.02 - 0.33)**	**(-0.97 - -0.27)**	**(-0.78 - -0.08)**	**(-0.97 - -0.26)**
Ventral[†]	0.24[†]	-0.023	-0.04	-0.03	-0.33[†]	-0.30[†]	-0.35[†]
	(-0.01 - 0.49)	(-0.12 - 0.08)	(-0.26 - 0.17)	(-0.12 - 0.07)	(-0.58 - -0.09)	(-0.54 - -0.06)	(-0.59 - -0.11)
Lateral	**0.60**	-0.12	**-0.51**	**-0.27**	-0.18	**-0.31**	**-0.39**
	(0.20 - 1.01)	(-0.32 - 0.09)	**(-0.95 - -0.08)**	**(-0.45 - -0.08)**	(-0.51 - 0.16)	**(-0.64 - 0.03)**	**(-0.72 - -0.05)**

[¥] Parameter estimates from linear mixed-effects model with confidence interval in parenthesis; $P \leq 0.05$ indicated in bold; intercept is the difference between face videos and baseline videos; motion is the proportion of scrubbed voxels; paradigm indicates which paradigm was used.

[œ] For this analysis, all infants younger than 5.0 months were selected (2.5-4.6 months, mean=3.6 months).

[†] The statistics for the Ventral parcel are circular because this set of subjects was chosen because they have the best face-selectivity in the Ventral parcel. Statistics are provided for illustrative reasons only.