Holistic face recognition is an emergent phenomenon of spatial integration in face-selective regions

Sonia Poltoratski[*,1], Kendrick Kay[2], Dawn Finzi[1], Kalanit Grill-Spector[1,3]

1. Department of Psychology, Stanford University

2. Center for Magnetic Resonance Research, University of Minnesota

3. Wu Tsai Neurosciences Institute, Stanford University

[*] corresponding author: sonia09@stanford.edu. Building 420, 450 Jane Stanford Way, Stanford, CA 94305.

1   Abstract

2   Spatial processing by receptive fields is a core property of the visual system. However, it is unknown

3   how spatial coding in high-level regions contributes to recognition behavior. As face inversion is

4   thought to disrupt typical 'holistic' processing of information in faces, we mapped population

5   receptive fields (pRFs) with upright and inverted faces in the human visual system. In face-selective

6   regions, but not primary visual cortex, pRFs and overall visual field coverage were smaller and shifted

7   downward in response to face inversion. From these measurements, we successfully predicted the

8   relative behavioral detriment of face inversion at different positions in the visual field. This

9   correspondence between neural measurements and behavior demonstrates how spatial integration

10   in face-selective regions enables holistic processing. These results not only show that spatial

11   processing in high-level visual regions is dynamically used towards recognition, but also suggest a

12   powerful approach for bridging neural computations by receptive fields to behavior.

13   Visual recognition is critical to human behavior, informing not only what we see, but also how

14   we navigate, socialize, and learn. In humans and non-human primates, visual recognition is carried

15   out by a series of hierarchical, interconnected cortical regions spanning from the occipital pole to the

16   ventral temporal cortex (VTC[1-3]). At the early cortical stages of this ventral visual stream, neurons

17   respond to spatially local regions in the visual field and are typically tuned to simple stimulus

18   properties like orientation or spatial frequency[4]. Later regions in the hierarchy show selectivity for

19   complex stimuli like faces or places, and their responses correspond to perceptual experience[5,6].

20   These responses are thought to be abstracted from and largely invariant to low-level image

21   properties like stimulus size or position in the visual field[7,8]. However, modern research suggests that

22   neural responses in VTC are sensitive to both stimulus size and position[9-14] providing a key empirical

23   challenge to classical theories. Yet, it remains unknown if and how spatial processing in high-level

24   visual regions is used toward recognition behavior.

25   The basic computational unit of sensory systems is the *receptive field* (RF), the region in

26   space to which a neuron responds[4]. Throughout the visual system, neurons with similar RFs in

27   retinotopic space are clustered in cortex, allowing researchers to use fMRI to measure the *population*

28   *receptive field* (pRF), or the portion of the visual field that is processed by the population of neurons

29   in a voxel[15]. Recently, our group and others have used pRF modeling to quantify spatial processing

30   in face-selective VTC regions, finding that pRFs in face-selective regions are progressively larger than

31   in early visual areas and are densely centered around the center of gaze (the fovea), resulting in

32   foveally-biased coverage of the visual field[16-18]. These response properties may enable neurons in

33   face-selective regions to integrate information across facial features for effective face recognition.

34   Indeed, pRFs and the resulting visual field coverage in face-selective regions become larger and

35   more foveal across childhood development, alongside with changes in fixation patterns and

36   improved recognition of faces[17]. These findings lead to the novel hypothesis that spatial processing

37   by pRFs in high-level visual regions may support recognition behavior.

38   To test this hypothesis, we examined the core human visual face network[19] (**Figure 1a**), here

39   defined as four anatomically distinct functional regions: one in the inferior occipital gyrus (IOG-

40   faces/OFA), one on the posterior fusiform gyrus (pFus-faces/FFA1), one straddling the mid-fusiform

41   sulcus (mFus-faces/FFA2), and one on the posterior end of the superior temporal sulcus (pSTS-

42   faces). This network serves as an ideal model system for deriving links between neural computations

43   and behavior, as it is reliably localized in each human participant, causally involved in face recognition

44   behavior[20,21], and stereotypically organized in relation to known anatomical landmarks and

45 cytoarchitectonic regions[22]. Importantly, face recognition behavior has been richly characterized
46 across decades of psychology research[23,24], allowing us to link neural responses to robust and
47 replicable behavioral markers of face recognition.

48   Broadly, we hypothesized that computations by pRFs in face-selective regions provide the
49 basis for spatial integration of information toward face recognition behavior. This led us to predict
50 that face inversion, which is thought to hinder typical spatial processing of faces[23,25-27], would alter
51 pRF estimates in face-selective regions, but not in early visual regions where spatial integration is not
52 sensitive to face content. To test our hypotheses, we mapped pRFs across the ventral visual pathway
53 with both upright and inverted faces (**Figure 1b**). We predicted that if large and foveal pRFs in face-
54 selective regions are adaptive for typical recognition behavior, both individual pRFs and the visual
55 field coverage afforded by a cortical region may be shifted in position and/or in size with face
56 inversion. Alternatively, if spatial coding in face-selective regions is a concomitant neural property
57 that does not actively play a role in face recognition, we would expect equivalent estimates when
58 mapping with upright or inverted faces in both face-selective regions and early visual cortex.

59   Finally, we hypothesized that if pRFs reflect neural mechanisms underlying functional face
60 recognition, they may link directly to the behavioral face inversion effect (FIE[23]. First described over
61 50 years ago, the FIE is a robust and specific behavioral deficit in recognizing faces when they are
62 presented upside down. Critically, the FIE has been associated with a failure of typical spatial pooling
63 of information across multiple features of a face[24,28], prompting us to examine the linkage between
64 the visual field coverage of face-selective regions and the behavioral FIE. We reasoned that if pRF
65 coverage in face-selective regions is reduced or shifted in response to face inversion, then neural
66 processing of inverted faces at the center of the visual field would be suboptimal and behavioral
67 recognition would be impaired. Accordingly, the behavioral FIE may be mitigated if faces are placed
68 in the optimal location for inverted faces as determined from the visual field coverage they elicit. As
69 we found that pRFs and the visual field coverage in face-selective regions where smaller and shifted
70 in response to inverted faces, we tested these predictions in a separate behavioral experiment
71 outside the scanner by using our participants' neural measurements to predict the magnitude of the
72 FIE across visual field positions.
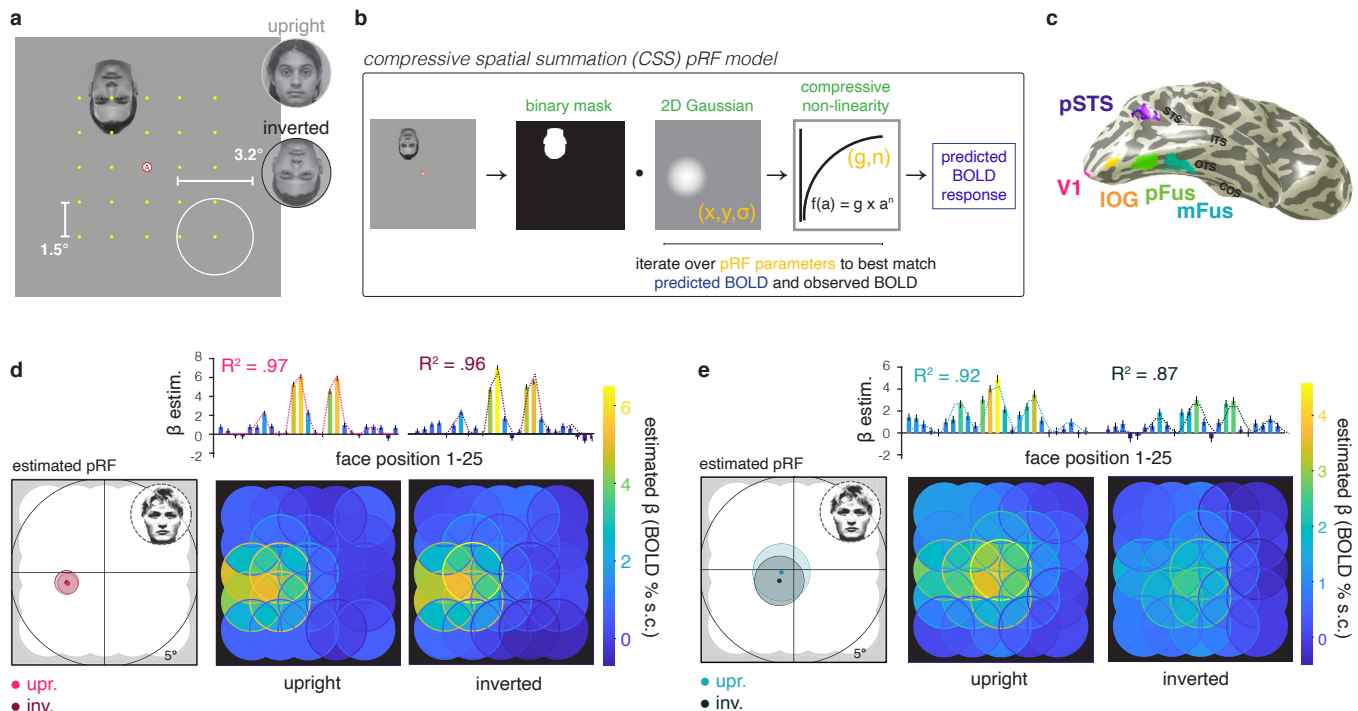
Figure 1. *Experimental design and pRF model fitting procedure.* **(a)** Schematic illustrating the 5 x 5 grid of positions (1.5° center-to-center) at which faces (3.2° diameter) were presented, either upright or inverted as shown. **(b)** Model fitting procedure of the compressive spatial summation (CSS[29]) pRF model, which represents the position of each face in the visual field as a binary mask. Estimated parameters (*yellow text*): center position X and Y, spread σ, exponent (n), and gain (g). While a single binary mask is shown here, an averaged silhouette mask of the faces that each participant saw at each position was used for the model fitting. **(c)** V1 and face-selective regions of interest on the cortical surface of a representative participant. *Acronyms:* V1: primary visual cortex; IOG: inferior occipital gyrus face-selective region, also referred to as the occipital face area (OFA); pFus: posterior fusiform face-selective region, also referred to as fusiform face area 1 (FFA-1); mFus: mid fusiform face-selective region, also referred to as fusiform face area 2 (FFA-2); pSTS: posterior superior temporal sulcus face-selective region. **(d, e)** Example V1 and mFus-faces voxel data and model fits. *Top (bar plots):* BOLD responses (β estimates in % signal change) for each of the 25 mapping positions for this voxel, plotted first from left to right, and then top to bottom. *Error bars:* ± SEM across runs; *Dotted lines:* pRF model fit, which was estimated separately for upright and inverted faces. *Left:* estimated pRF in response to upright (*light colors*) and inverted (*dark colors*) faces. Circles contours indicate estimated pRF size, radius = σ / √n. *Right (overlapping-disk) panels:* Responses at each position at which the mapping stimuli appeared, colored by response amplitude. Overlapping stimuli responses are averaged within each disc, while the absolute responses at each position are indicated by the disc outline.

Results

73

74 To map population receptive fields in response to upright and inverted faces, twelve

75 participants took part in an fMRI experiment. Faces were presented in randomized order at 25

76 locations spanning the central ~9.2° of the visual field while participants maintained fixation on a

77 central letter stimulus (**Figure 1a**; Online Methods); fixation is critical for these measurements to

78 ensure that the retinotopic position of the stimulus can be reliably coded in the model. During pRF

79 mapping, participants performed a challenging rapid visual stimulus presentation (RSVP) letter task

80 at fixation. While they were aware that faces would appear on the display, participants were

81 instructed that faces were not task-relevant during the fMRI session. We saw no differences in

82   behavioral performance on the letter task between trials in which upright (90.2 ± 1.8% correct) or

83   inverted (89.4 ± 2.5%) faces were presented ($t(11) = 0.531$, p = 0.606). Therefore, we infer that any

84   observed differences in pRFs measured in response to upright and inverted faces are stimulus-driven

85   by the face inversion, rather than by task performance or attention.

86   We implemented the compressive spatial summation (CSS) pRF model[16,29], which models

87   each voxel's pRF as a circular Gaussian with a compressive nonlinearity (n; **Figure 1b**). Each model

88   fit yields five parameters that describe the gain (g), position (x,y), and size ($\sigma/\sqrt{n}$) of the pRF of each

89   voxel, as well as a measure of the goodness-of-fit of the model (R^2). pRFs were estimated

90   independently for upright and inverted faces in each voxel in the primary visual cortex (V1) and four

91   face-selective cortical regions (**Figure 1c**); data from V1 is presented as an early visual control. As

92   illustrated for example voxels from V1 (**Figure 1d**) and mFus-faces (**Figure 1e**), the position of the face

93   in the visual field robustly modulates the response of these regions. Further, the CSS pRF model

94   accurately quantifies these voxels' responses, as it explains the majority of their respective variances

95   (**Figure 1d,e-R^2**). Notably, the independent model fits yield similar pRF estimates across upright and

96   inverted faces in V1, corresponding to predictions that spatial responses in early visual cortex are

97   largely agnostic to this manipulation of stimulus content (**Fig 1d-left panel**). However, for the example

98   mFus-faces voxel, face inversion yields responses that are lower in magnitude and shifted in the

99   visual field, which consequently produce differential pRF estimates across the two mapping

100   conditions (**Fig 1e-left panel**).

101

102   *Mapping with inverted faces modulates pRF properties in face-selective regions but not V1*

103   We first quantified and compared average pRF properties for upright and inverted faces in

104   face-selective regions. We found no significant differences in the horizontal position of pRF centers

105   across upright and inverted mapping conditions in face-selective regions (**Figure 2a, Table S1**).

106   However, face inversion yielded pRF centers that were consistently shifted downwards (**Figure 2b**).

107   We quantified changes in estimated pRF parameters using a 3-way repeated measures analysis of

108   variance (ANOVA) with factors of ROI (IOG-/pFus-/mFus-/pSTS-faces), hemisphere (right/left), and

109   mapping condition (upright/inverted). This revealed a significant effect of inversion on pRF center Y

110   position ($F(1,10) = 8.82$, p = 0.014; no other significant main effects or interactions, **Table S1**). Post-

111   hoc t-tests revealed that downward shifts in the average pRF position were significant in all ventral

112   face regions (IOG-, pFus-, mFus-faces, **Figure 2b**). Additionally, pRF gain was significantly lower for

113   inverted than upright faces in some of the face ROIs (**Figure 2c**; significant ROI by inversion

114   interaction F(3,30) = 4.66, p = 0.009, 3-way repeated-measures-ANOVA; no other significant effects,

115   **Table S1**). Post-hoc t-tests showed that pRF gains were lower for inverted than upright faces in IOG-

116   and mFus-faces (**Figure 2c**). This finding is consistent with prior reports of lower mean responses for

117   inverted than upright faces in ventral face regions[25,26,30].
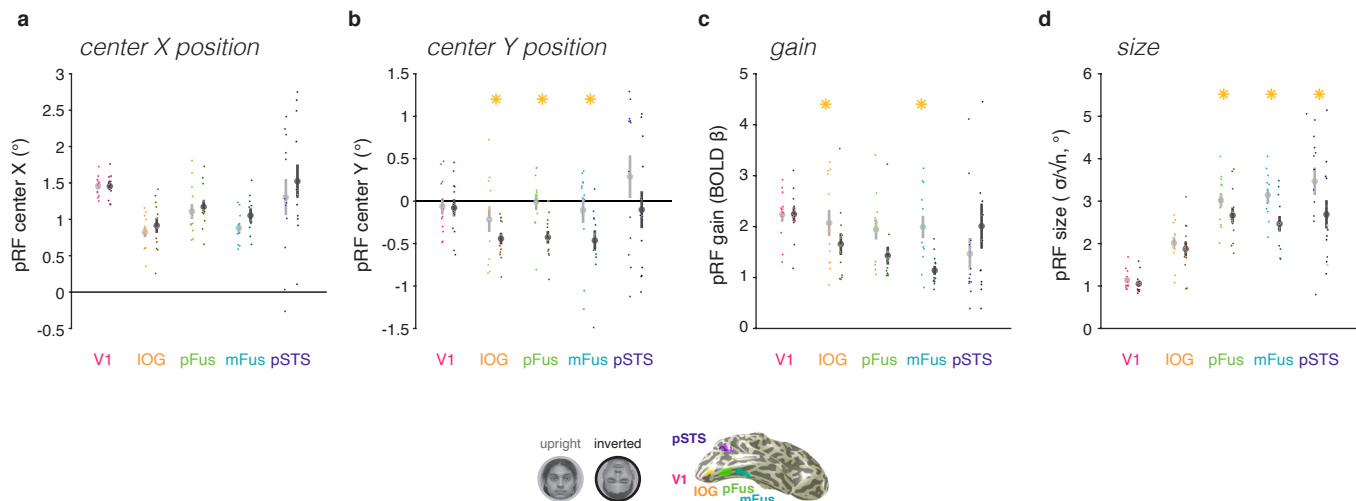
Figure 2. *Face inversion impacts pRF estimates in face-selective cortical areas but not primary visual cortex.* Mean pRF properties when mapped with upright faces (*light grey)* and inverted faces (*dark grey*) in bilateral regions of interest across 12 participants. *Error bars:* ± SEM across participants. *Dots*: individual participant means. *Asterisks:* statistically significant differences between upright and inverted face mapping conditions (two-tailed paired t-test across participants, p < .05). **(a)** pRF center X position; right hemisphere data are mirror flipped to left. **(b)** pRF center Y position. **(c)** pRF gain. **(d)** pRF size (σ / √n).

118   Interestingly, there were also significant differences in estimated pRF size across upright and

119   inverted mapping conditions, which varied by face-selective region (ROI by inversion interaction

120   F(3,30) = 7.53, p = 0.0001, 3-way repeated-measures-ANOVA, **Table S1**). Specifically, pRF size was

121   significantly smaller in fusiform face-selective regions (pFus- and mFus-faces) and pSTS-faces in

122   response to face inversion (*t'*s(11) > 2.98, *p*'s < 0.013; **Figure 2d**), but not in IOG-faces.

123   Overall, face inversion had the most pronounced effect on pRF properties in mFus-faces, the

124   end stage of the ventral face-processing hierarchy[18,19]. In mFus-faces, face inversion yielded a

125   significant downward shift of pRFs, smaller pRF size, and smaller gain. As we did not find any

126   significant difference across hemispheres of the effect of inversion on pRF responses (**Table S1**), we

127   combine data across hemispheres unless noted. In contrast to the profound effect face inversion

128   had on pRF properties in face-selective regions, we saw no significant differences between mapping

129   conditions in any estimated pRF property in V1 (**Figure 2-V1**). This suggests that spatial integration

130   is altered in response to face inversion specifically in higher-level face-selective regions, but not in

131   early visual cortex.

132

133   *Signal-strength differences are insufficient to explain the observed effects*

134   As BOLD amplitudes were generally reduced in face-selective regions in response to inverted

135   faces, we observed lower estimated pRF gain (**Figure 2c**). Additionally, model goodness-of-fit in face-

136   selective regions was consistently lower in the inverted than in the upright mapping condition (**Figure**

137   **S1**). We asked if these reductions in signal strength might explain the observed differences in pRF

138   position and size. To test this possibility, we simulated the effect of changing the magnitude of

139   responses and level of noise in mFus-faces on model estimates of pRF center Y position and size.

140   Results of the simulations suggest that differences in signal strength differences are insufficient to

141   account for the scale of our observed effects in mFus-faces (**Figure S2b-c**).
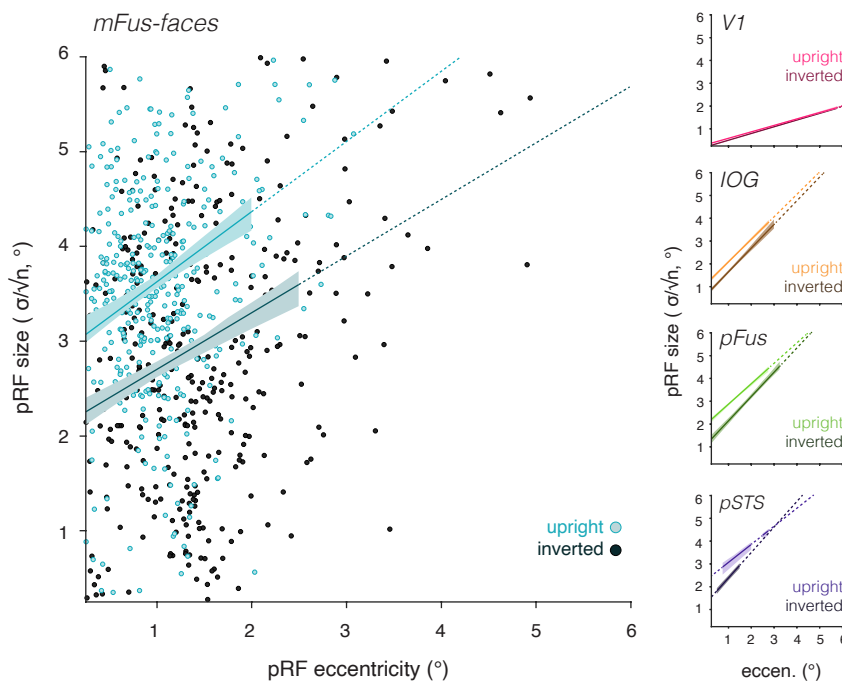
Figure 3. *Face inversion yields smaller pRFs in face-selective regions across eccentricities.* The relationship between size and eccentricity in each ROI, for voxels mapped with upright faces (light colors) and inverted faces (dark colors). This relationship is summarized by the line of best-fit across all voxels pooled over participants with model goodness of fit R^2 > .5. *Shaded error:* a bootstrapped 68% confidence interval at bins (0.25° wide) for which there were more than 20 voxels across participants. *Dashed lines:* bins for which there were fewer than 20 voxels across participants. *Dots:* a random sample of 300 mFus-faces voxels across participants. See also **Figure S3**.

142   *pRFs in face-selective regions are smaller across eccentricities in response to inverted faces*

143   The analyses of mean pRF size are consistent with the hypothesis that face inversion yields

144   more spatially constrained integration of information via pRFs in face-selective regions. However, as

145   pRF size linearly increases with eccentricity throughout the visual hierarchy[15,16], it is critical to

146   consider the relationship between size and eccentricity when evaluating pRF size estimates across

147   conditions. Thus, we compared the size vs. eccentricity relationship for upright and inverted faces in

148   each ROI. We reasoned that if there are differences in spatial processing across mapping

149   conditions,we should observe a higher intercept or a steeper slope relating size to eccentricity for

150   pRFs mapped with upright than with inverted faces.

151   As evident in **Figure 3**, we observed smaller pRF sizes in response to inverted than to upright

152   faces across eccentricities in pFus- and mFus-faces. In contrast, we saw minimal differences in pRF

153   sizes across eccentricities between upright and inverted face mapping for either IOG-faces or V1,

154   consistent with data in **Figure 2d**. To evaluate these results across participants, we performed an

155   equivalent line-fitting separately for each participant (**Figure S3**) and tested for significance. Results

156   indicate the slope relating pRF size and eccentricity is not affected by face inversion (F(1) = 0.04, p

157   = 0.852 , 2-way repeated measures ANOVA on slope with factors of ROI (IOG-/pFus-/mFus-/pSTS-

158   faces) and condition (upright/inverted; **Table S2**). However, pRFs were smaller across eccentricities

159   (significant main effect of intercept (F(1) = 9.08, p = 0.012, 2-way repeated measures ANOVA, **Table**

160   **S2**). Post-hoc t-tests revealed that these effects were significant in both pFus- and mFus-faces

161   (pFus: t(11) = 4.185, p = 0.002; mFus: t(11) = 2.371, p = 0.037) but not in IOG-faces, pSTS-faces,

162   or V1.
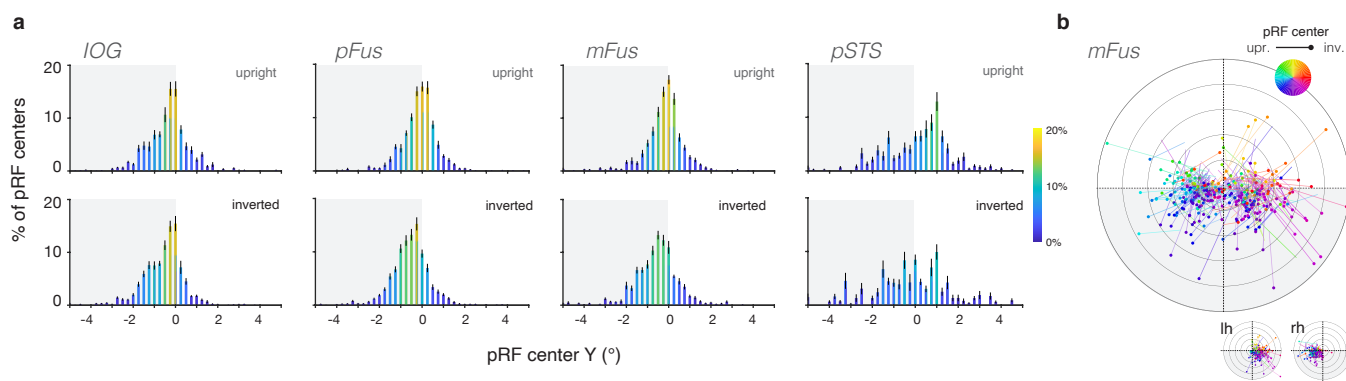
Figure 4. *In face-selective regions, inverted-mapped pRFs differ in vertical position Y.* **(a)** Across participants, the average distribution of pRF position Y when mapped with upright (top row) and inverted (bottom row) faces. *Error bars:* ± SEM across participants in each 0.25° bin. **(b)** Position shifts for a random subsample of 300 mFus-faces voxels across participants (for which the model variance explained is higher than 50% of the variance); voxel pRF center position estimates are consistently shifted downward.

163   *pRFs in face-selective regions are shifted downward in response to inverted faces*

164   To further explicate the effect of face inversion on pRF position in face-selective regions, we

165   calculated the distributions of pRF positions in each participant and ROI when mapped with upright

166   and inverted faces (**Figure 4**). In IOG-faces, the average distribution of pRF center Y positions

167   mapped with upright (**top panel**) and inverted (**bottom panel**) faces were largely equivalent (**Figure**

168   **4a**). In contrast, face inversion yielded a robust downward shift in pFus-, mFus-, and pSTS-faces.

169   In fusiform face-selective regions, inversion shifted the overall distributions' centers from the

170   horizontal meridian to the lower visual field. These shifts are visible when plotting the location of

171   each voxel's upright vs. inverted pRF position as shown in **Figure 4b** for mFus-faces. While we also

172   see some variability in the estimated pRF center X position across upright and inverted mapping,

173   these shifts have no consistent direction, even when considering hemispheres independently

174   (**Table S1, Figure 2a**). The downward shift in pRF center position in response to inverted faces

175   altered the overall distribution of pRFs to be further from the fovea. Notably, increased

176   eccentricities would have predicted larger pRFs mapped with inverted than with upright faces,

177   which is the opposite of our empirical findings (**Figure 3**).

178

179   *pRF changes result in differences in the coverage of the visual field in face-selective regions*

180   Given our findings that pRFs in face-selective regions are modulated by face inversion, we next

181   asked how these changes may impact the overall coverage of the visual field afforded by these

182   regions. This visual field coverage reflects the distributed spatial responsivity of the entire population

183   of neurons a cortical region, and thus may be more closely linked to the behaviors that the population

184   subserves. Thus, we constructed and quantified density coverage maps in each ROI for each

185   participant and then averaged across participants (see Online Methods, **Figure 5**).

186   When pRFs are mapped with upright faces, ventral face-selective regions show a prominent

187   foveal bias in coverage (**Figure 5-left-U**). This foveal bias is markedly reduced in response to inverted

188   faces (**Figure 5-left-I**), and is coupled with a significant reduction in the coverage area and downwards

189   shift of the coverage (**Table S3, Figure S4**). These effects appear largely consistent across individual

190   hemispheres (**Figure S5**). As in voxelwise analyses, mFus-faces shows the strongest differentiation

191   between upright- and inverted-mapped coverage (**Figure 5**). Coverage for the latter spans a smaller

192   portion of the visual field (significant change in full-width-half-max area: $t(11) = 2.79$, p = 0.017) and

193   is shifted downward relative to the former (significant change in center-of-mass Y: $t(11) = 3.338$, p =

194   0.007, but not center-of-mass X: $t(11) = -0.569$, p = 0.58, **Figure S4**). A similar downward shift is

195   evident in pFus-faces (significant change in center-of-mass Y: $t(11) = 3.123$, p = 0.01). In contrast,

196   IOG-faces, pSTS-faces and V1 showed minimal differences in coverage between the mapping

197   conditions.
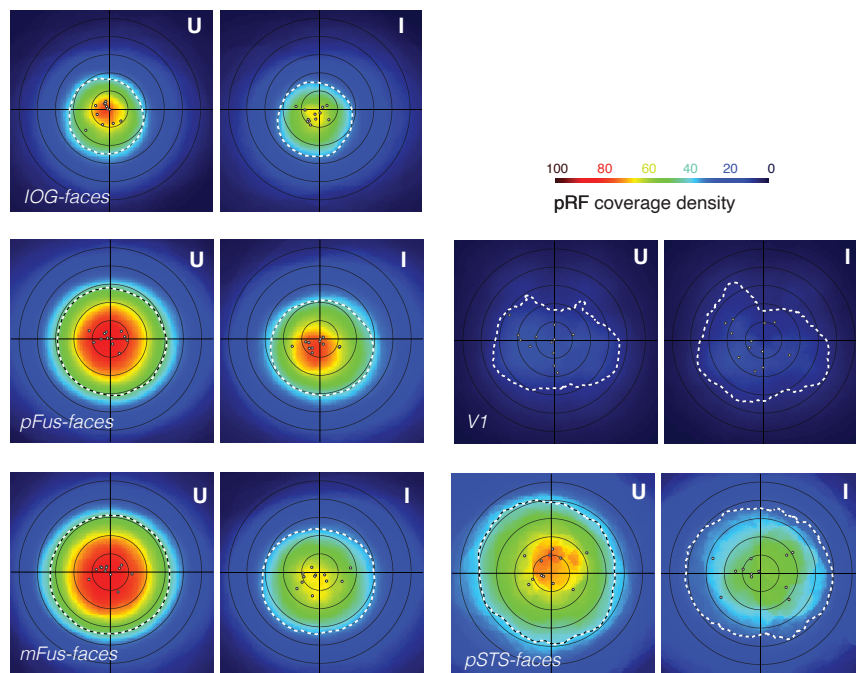
Figure 5. *Face inversion leads to differences in overall visual field coverage in face-selective regions.* **(a)** Average pRF density plots for each region across participants (N = 12) when mapped with upright (left) and inverted (right) faces. Density plots are bootstrapped for each participant using 1000 draws of 80% of the voxels and then averaged across participants (see Online Methods). The resulting plot color indicates the average percentage of voxels whose pRFs cover each point in retinotopic space, where coverage is defined as within $2\sigma / \sqrt{n}$ ° from the pRF center . *Dashed white contours:* full width half maximum (FWHM) of the coverage density, which we use as a summary of the overall coverage field of view of the region. *Dots*: coverage center of mass of each individual participant.

198   Given prior reports of stimulus-dependence of visual field coverage in nearby word-selective

199   regions[31], we sought to examine whether differences in coverage in face-selective regions were

200   specific to inversion. We evaluated this by measuring coverage in our participants' face-selective

201   regions to different stimuli: brightly colored cartoon images in a separate pRF mapping experiment

202   (**Figure S5a,** [18]). Coverage in face-selective regions mapped with these highly varied stimuli was

203   similar to the upright faces condition; pRF density was highest around the fovea, and consequently

204   the coverage was centered on the fovea (**Figure S5b-c**). This suggests that reduction in area and the

205   downward shift of coverage observed for inverted faces is specific to face inversion, and does not

206   occur for any suboptimal mapping stimulus.

207   In all, we see profound differences in the way that face-selective regions sample visual space in

208   response to face inversion, both at the level of individual voxels, and by the collection of voxels

209   constituting a region. Importantly, pRF differences are observed even in response to passively-

210   viewed mapping faces, as participants' attention was diverted to an unrelated RSVP task on letters

211   at the central fixation. This is particularly interesting as it suggests that the reported effects of

212   inversion on spatial processing in face-selective regions is not a simple reflection of the participants'

213   task or focus of attention, but instead reflect an underlying sensory property that may ultimately

214   constrain behavior. To evaluate this idea, we next used our fMRI measurements to derive and test a

215   specific prediction for participants' performance in recognizing upright and inverted faces outside of
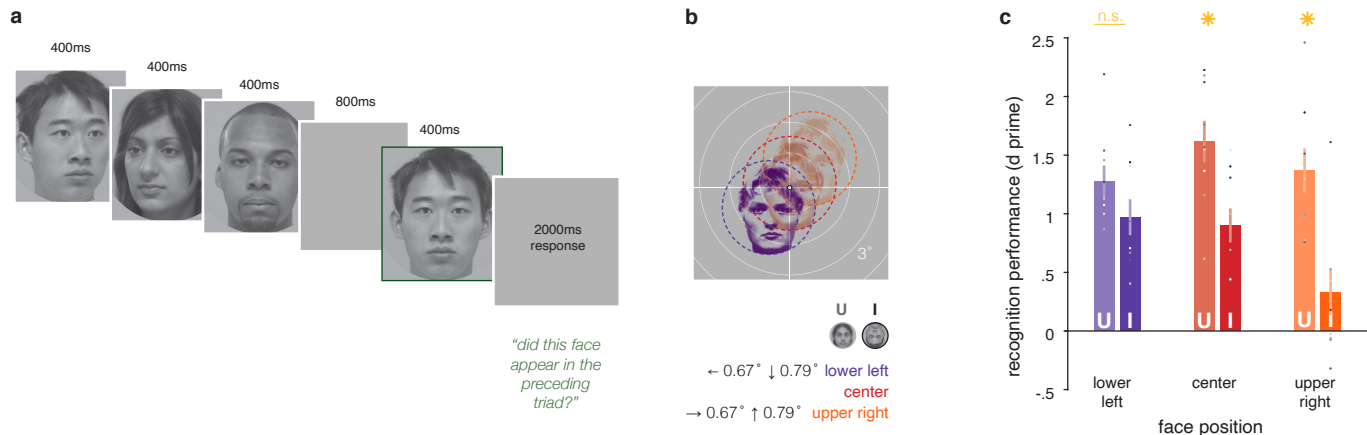
216   the scanner.
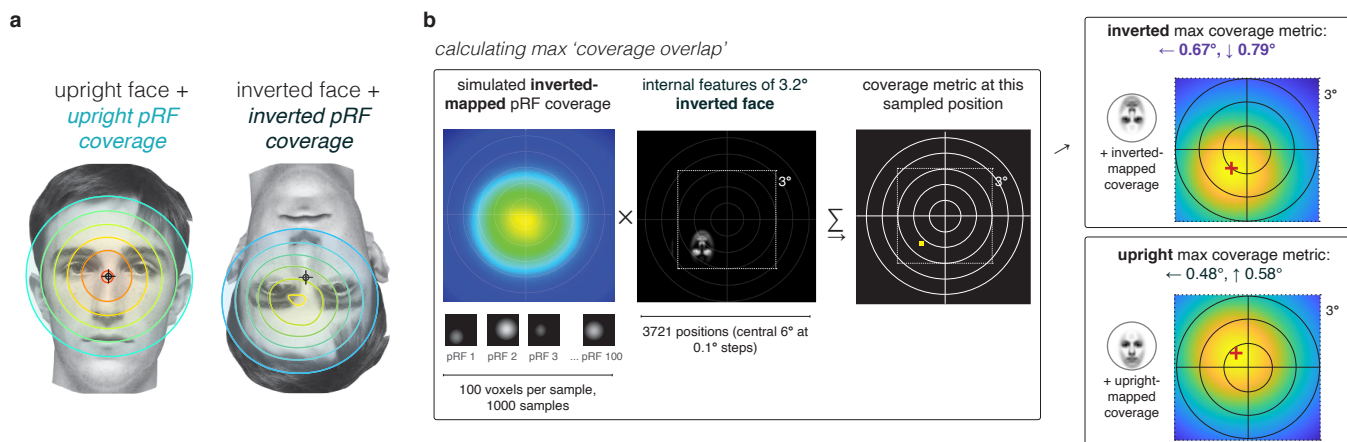
Figure 6. *Using visual field coverage of face-selective regions to predict the location of improved performance for inverted faces.* **(a)** Schematized depiction of upright- and inverted-mapped mFus-faces pRF coverage when one views example faces centered on fixation. This predicts that viewing inverted faces yields suboptimal coverage of the location of the face features by pRFs in face-selective regions, which may contribute to the reported behavioral detriments in recognizing inverted faces (e.g. the face inversion effect, or FIE). **(b)** To evaluate the hypothesis schematized in (a), we used a simulation procedure (see Online Methods) to determine the location in the visual field that would produce maximal pRF overlap with the internal features of 3.2° faces. To do so, we iteratively sampled a subset of mFus-faces voxels to generate bootstrapped coverage in response to inverted or upright faces; we then took the dot product of this bootstrapped coverage and a face centered at various retinotopic positions (central 3° x 3°, sampled at 0.1° steps). The retinotopic location that yielded the maximal overlap (dot product) between the inverted face and inverted-mapped pRF coverage, as well as the upright face and upright-mapped coverage, is shown in the right panels.

A novel test of the relationship between pRF coverage and face recognition behavior

218   During typical face viewing, pRF coverage in face-selective areas seems optimized to capture

219   information from the face: large and foveal pRFs extend into the contralateral visual field, allowing for

220   the processing of many features of a centrally-viewed face by the pRFs of a face-selective region[16].

221   However, as schematized in **Figure 6a**, our data suggest that this is not the case when viewing

222   inverted faces. If participants fixate at the center of an inverted face in a standard recognition task,

223   our data predict that the visual field coverage by pRFs of face-selective regions involved in

224   recognition will suboptimally cover the inverted facial features (**Figure 6a**). This leads to an intriguing

225   hypothesis about the functional role of spatial integration in face-selective regions: if pRF coverage

226   corresponds to windows of spatial integration toward face recognition, then increasing the overlap

227   of this coverage with inverted facial features should yield better recognition performance of inverted

228   faces, consequently reducing the magnitude of the behavioral face inversion effect (FIE) in that
229   location in the visual field.

230   To evaluate this hypothesis, we first used a simulation to determine the location at which we
231   expected the behavioral FIE to be attenuated (**Figure 6b**, Online Methods). As a large body of
232   behavioral research indicates that internal facial features are key for recognition[23,25-27], we reasoned
233   that performance for inverted faces would be best at the retinotopic location at which its internal
234   features maximally overlap with the inverted-mapped pRF coverage in face-selective regions. To
235   determine this location, we simulated the overlap between the internal features of inverted and
236   upright faces and the corresponding pRF coverage from mFus-faces, where the effects of face
237   inversion were strongest. As in the fMRI study, we used 3.2° face images in the simulation, and
238   determined that the optimal location to place these inverted faces was centered 0.67° leftward and
239   0.79° downward from fixation. Notably, this location is different than the optimal location for upright
240   faces (**Figure 6b-bottom right**). While in both cases, predicted performance is best in the left visual
241   field[32,33], only performance for inverted faces is predicted to be improved in the lower visual field (see
242   also **Figure S7**).

Figure 7. *The behavioral FIE is reduced at the retinotopic position that maximally overlaps mFus-faces pRF coverage.* **(a)** Illustration of the trial structure and timing of the recognition task outside the scanner, in which participants viewed a triad of rapidly-presented faces and indicated whether the subsequent target face had appeared on the trial. Within a trial, faces could be upright (as illustrated), or inverted, and appeared at the three locations described in (b). Subjects fixated on a central bullseye target (0.3° diameter), and their fixation was monitored with an eyetracker. Only trials on which participants maintained fixation were included in analyses. **(b)** Illustration of the three positions at which faces (3.2° diameter) appeared on the screen, as determined by mFus-faces pRF coverage from the fMRI experiment (**Figure 6b**). **(c)** Results of the behavioral experiment. When participants (N=9) viewed upright faces (light color bars), performance was predictably highest at the center (red), and fell when the faces were presented in the lower left (purple) or upper right (orange) positions. However, recognition performance for inverted faces (dark color bars) was highest at the lower-left location. In fact, across participants, the behavioral FIE (upright minus inverted performance) was significantly reduced in the lower-left position, while remaining strong at the center and upper-right positions. *Dots:* individual participant performance; *Asterisks:* significant differences between upright and inverted face conditions at p < .05. *N.S.:* not significant.

243   We tested these predictions by evaluating face recognition performance of each of our

244   participants in a behavioral experiment conducted outside of the scanner. The behavioral study

245   consisted of a challenging recognition memory task on upright and inverted faces (50% each, **Figure**

246   **7a**). Fixation was held at a small central bullseye, and strictly monitored (Online Methods). On each

247   trial participants saw a sequence of 3 faces (all upright or all inverted), each shown for 400ms (**Figure**

248   **7a**). After an 800ms interval, a target face was shown, and participants were asked to indicate

249   whether the target was present in the preceding triad (50% probability). Critically, to test our

250   hypothesis, recognition performance on upright and inverted faces was evaluated at three retinotopic

251   locations: the *lower left*, corresponding to maximal pRF overlap for inverted faces, *central* fixation,

252   and an *upper right* location that was equivalent in distance from fixation (**Figure 7b**). This third

253   condition serves as a control; if retinal acuity alone determines performance, recognition should

254   worsen symmetrically with increasing distance from the fovea for both upright and inverted faces.

255   Notably, these position shifts are relatively small: 3.2° faces are shifted 1.03° off-center in the lower-

256   left and upper-right conditions. Individual trials on which participants broke fixation were excluded

257   from the analysis, and the data of three participants who were unable to maintain consistent fixation

258   were removed (see Online Methods).

259   We found striking differences in recognition performance at the three probe locations.

260   Recognition for upright faces was best when faces were presented at the center of the screen and

261   declined similarly when placed at the lower left or upper right positions (**Figure 7c-U**). This is

262   consistent with our predictions, as both the upper-right and lower-left positions were nearly

263   equidistant from the simulated optimal position for upright faces (**Figure 6b;** the lower left is 0.22°

264   farther). Recognition of inverted faces, however, followed a different pattern: performance was equal

265   if not better at the lower-left location than at fixation, but starkly worse at the equidistant upper-right

266   location (**Figure 7c-I**). Indeed, the magnitude of the FIE, or the difference between recognition

267   performance for inverted and upright faces, was not significantly different from zero at the lower-left

268   location ($t(8) = 1.68$, $p = 0.13$, 2-sided paired t-test), while remaining significant at the other locations

269   (center: $t(8) = 3.43$, $p = 0.01$, upper right: $t(8) = 5.06$, $p < 0.01$). Thus, the behavioral FIE is attenuated

270   by placing faces in the location where inverted-mapped coverage of the visual field in ventral face-

271   selective regions covers facial features optimally, linking the effect of face inversion on visual field

272   coverage in face-selective regions to face recognition performance.

273   Discussion

274   Our experiments demonstrate stimulus-driven changes in spatial coding of face-selective
275   regions that correspond directly to visual recognition behavior. In the neuroimaging experiment, we
276   demonstrated that pRFs and coverage of the visual field in the face network, but not in primary visual
277   cortex, are altered by face inversion. While each of the face-selective regions showed some degree
278   of sensitivity to inverted faces, effects were most pronounced in mFus-faces (FFA-2), where both
279   individual pRFs and the coverage of the visual field were shifted downward and reduced in response
280   to inverted faces. A similar pattern of results was found in pFus-faces (FFA-1). pSTS-faces, a region
281   in the lateral stream of the face-selective VTC network thought to process dynamic aspects of
282   faces[34,35], exhibited similar effects of smaller pRF size, albeit noisier responses overall. This may be
283   attributed to the substantially more peripheral responses in pSTS-faces[18,36], making our mapping
284   with still images over a 9.2° x 9.2° visual field suboptimal. However, face inversion did not significantly
285   alter visual field coverage in IOG-faces. This is consistent with the region's strong responses to
286   individual face features[14,37], such as the eyes or mouth, on which inversion has a smaller perceptual
287   effect. Importantly, we found differences in spatial processing in response to face inversion while
288   participants performed a task that removed their directed attention from the faces. Following prior
289   behavioral findings[38], this suggests that bottom-up properties of the stimulus itself, and not attention
290   or task drive the modulation of pRF properties by face inversion.

291   In the behavioral experiment, we demonstrated that the magnitude of the face inversion effect
292   (FIE) varies across the visual field in a manner that is strikingly consistent with our measurements of
293   differential visual field coverage in the fMRI study. Specifically, we showed that the magnitude of FIE
294   can be mitigated by placing faces in an optimal location that corresponds to the inverted-mapped
295   visual field coverage of mFus-faces. These data suggest that the FIE is a consequence of the
296   differential population activity in face-selective regions evoked by face inversion, which corresponds
297   to large-scale differences in the way these regions sample the visual field. More generally, our results
298   highlight a precise relationship between spatial processing in high-level visual regions and the
299   recognition behavior that they subserve.

300   *Holistic face processing as spatial integration across multiple features*

301   An extensive body of behavioral research has shown that faces are processed holistically—
302   that is, face recognition behavior involves concurrent processing of the whole face, rather than its

303   individual features[23,27,28,38-42]. However, despite decades of research into the behavioral signatures of

304   holistic processing of faces, its underlying neural mechanisms remain opaque.

305   From our findings, we propose that holistic processing of faces can be understood as a

306   consequence of spatial integration of information across face features, which emerges from the large

307   and foveal receptive fields of ventral face-selective regions. Consequently, behavioral failures of

308   typical holistic processing, like the FIE[23], appear to be a result of maladaptive, stimulus-driven

309   changes of pRF coverage in these regions. These findings provide neural support to cognitive

310   theories that predicted a reduced perceptual field for inverted faces[28]. Importantly, our data not only

311   elucidate the neural and computational mechanisms of face inversion, but also provide for the first

312   time a computational framework that generates quantitative predictions of behavioral deficits of

313   holistic processing. This provides an important foundation for future work that can quantitively

314   examine if individual differences in holistic processing[43] and impairments in face recognition[44,45] are

315   derivative of differences in pRF size, position, and coverage in high-level visual regions.

316

317   *Stimulus selectivity interacts with spatial processing by pRFs*

318   Our data demonstrate an interaction between estimates of pRF properties in face-selective

319   regions and the stimulus used in mapping (upright vs. inverted faces), which is specific to face

320   inversion as we did not find equivalent deviations for phase-scrambled faces[16] or diverse cartoons[18]

321   **Figure S6**).

322   What aspect of face inversion produces the observed changes in pRFs? One possibility is

323   that the differential location of the eyes or internal features contributes to pRF differences with face

324   inversion. In our stimuli, the direction and magnitude of the observed Y position shift in pRF centers

325   appears to correspond to the differential center-of-mass of the internal features between upright and

326   inverted faces rather than solely the eyes (**Figure S8**). However, the differential location of internal

327   features is not sufficient to explain the observed reduction in pRF size nor the lower gain. Thus, an

328   interesting direction for future research is building an image-computable model that predicts all

329   measured effects of face inversion. To do so, researchers may consider how the displacement of

330   internal features may interact with the learned spatial configuration of typically-viewed, upright

331   faces[46], potentially yielding suboptimal contextual modulation of RFs of individual neurons[47].

332   *High-level pRF processing as a bridge from brain to behavior*

333   While spatial sensitivity has been extensively demonstrated in VTC[9-14], the prevailing view

334   theorizes that spatial information in ventral temporal regions is independent from the recognition that

335   these regions support. Our findings emphasize that high-level visual regions not only maintain

336   spatially localized processing that can be computationally modeled by pRFs, but also that this

337   processing dynamically supports visual recognition behavior. Doing so provides the first

338   demonstration of the functional utility of spatial processing in the ventral 'what' stream.

339   More broadly, as receptive fields (RFs) are a basic computational unit of many domains[48-52],

340   we propose that quantitative measurement and modeling of pRFs in high-level regions across the

341   brain opens compelling avenues for deriving precise links between brain and behavior – especially

342   as activity in these regions often corresponds to perceptual experience more directly than in earlier

343   stages of neural processing. In particular, our approach can not only quantitatively determine what

344   factors may alter pRFs in other domains, but also leverage these measurements to generate new

345   testable behavioral predictions.

346   In sum, this work demonstrates a precise link between spatial processing in face-selective

347   regions and recognition behavior. We find that the neurally ill-defined concept of 'holistic' face

348   processing is an emergent property of spatial integration by population receptive fields. These data

349   suggest promising new directions for bridging basic neural computations by receptive fields to

350   behavior.

351

352   Methods

353   Participants: 13 participants (6 women) ages 20-31 participated in a single session of pRF mapping,

354   including authors SP and DF. Nine participants identified as white/Caucasian, 1 as Asian, 1 as Black,

355   and 2 as mixed-race (Hispanic/white and Asian/white). One participant's data was excluded for

356   excessive motion (> 4mm) during the scan. All participants provided informed, written consent per

357   the Internal Review Board of Stanford University, and were compensated for their time. No statistical

358   modeling was used to determine the size of our participant sample as we model each voxel

359   individually in each participant. We used typical sample sizes for fMRI pRF studies of the visual

360   system which use data from a range of 3-20 participants[15,16,31,53-56]. All scan participants completed

361   the subsequent behavioral experiment.

362   <u>Scanning protocol:</u> Participants underwent a single fMRI session in the main experiment, which

363   consisted of 8-10 runs of pRF mapping and lasted approximately 90 minutes. Scanning was done

364   at the 3 Tesla GE research-dedicated magnet at the Stanford Center for Cognitive and

365   Neurobiological Imaging (CNI). Functional T2* weighted pRF mapping runs were each 282s long,

366   and collected using a 32-channel headcoil, single-shot EPI, with a voxel resolution of 2.4mm

367   isotropic and a TR of 2s (FOV = 192, TE = 30ms, flip angle = 77°). Twenty-eight slices were

368   prescribed parallel to the parieto-occipital sulcus to cover each participant's occipital and ventral

369   temporal lobes.

370   We additionally acquired a high-resolution (1mm isotropic voxels) T1-weighted full-brain anatomical

371   during either the main experimental session or in a separate retinotopy/localizer session for each

372   participant, and used this whole-brain anatomical image to align data across the experiment,

373   localizer, and retinotopy sessions in native participant space.

374   <u>pRF mapping</u>: Each population receptive field mapping run lasted 282 seconds, and presented face

375   images at 25 spatial positions to span the visual field over time (**Figure 1a**). Face images were 3.2°

376   in diameter, and were offset 1.5° center-to-center to tile approximately 9.2° x 9.2° of the central

377   visual field. Faces were presented in 4s trials of faces (50 trials, 25 positions x upright/inverted) and

378   blank periods (10 trials), which were randomly ordered for each run and each participant. Each run

379   began and ended with an additional 16s blank period. Throughout each run, participants performed

380   a challenging RSVP letter detection task at the central fixation point (0.4° diameter). Participants'

381   fixation was monitored with an Eyelink 1000 eyetracker during the scan.

382   *Trial structure:* Each trial consisted of 7 faces (300ms on, 200ms off) and one blank interval (500ms)

383   presented at a single spatial location. The same timing structure was used for the central letter

384   presentation, including during no-face blank periods.

385   *RSVP task:* On each trial, participants were instructed to attend and respond to a stream of small,

386   rapidly presented letters (A-S-D-F-G-H-J-K) and detect one-back repeats of the same letter. Half of

387   all trials contained a repeat in the 7-letter stream, which occurred randomly in time within the trial.

388   Letters were presented throughout the entire mapping run, and participants reported the task to be

389   both difficult and attentionally engaging. Overall performance ranged from 77% to 98% correct

390   (mean = 90%, SE = 2%), and we observed no difference between performance on trials that

391   contained upright vs. inverted faces ($t(11) = 0.531$, $p = 0.606$), confirming that participants' attention

392   was not systematically drawn to either upright or inverted faces.

393   *Face stimuli:* We used front-view, grey-scaled face photographs previously used in Kay et al.[16] and

394   other work in the lab (VPNL-101-faces). These depicted 95 different individuals of various genders

395   and races in a neutral expression; demographics represented that of the Stanford University

396   undergraduate student body. Faces were presented in a circular aperture and were matched on the

397   overall height of the head, resulting in substantial consistency in the position of internal facial features

398   and some variability in the position and presence of external features, including hair.

399   ROI definition: Each participant also completed an fMRI session of retinotopic mapping (4 runs) using

400   brightly colored, 8Hz cartoon frames in a traveling wave aperture[18], as well as 3-10 runs of a visual

401   category localizer[57]. The localizer data was used to define face-selective regions in each participant

402   by contrasting responses to images of child faces + adult faces versus images from eight other

403   categories including bodies, limbs, objects, places, words and numbers, with a threshold of $t \geq 2.7$,

404   voxel level. Retinotopic regions V1-hV4 were defined from the retinotopy scans; boundaries between

405   retinotopic areas were delineated by hand by identifying reversals in the phase of the polar angle

406   map measurements and anatomical landmarks, as detailed in Witthoft et al.[58]. Face-selective region

407   IOG was defined to exclude any voxels overlapping hV4.

408   Data preprocessing: fMRI data was preprocessed following a standard pipeline using FSL

409   (https://fsl.fmrib.ox.ac.uk/fsl/fslwiki/Fslutils) and mrVista (https://github.com/vistalab) tools. Data

410   from the experimental runs and the ROI masks derived from the retinotopy and localizer analyses

411   were co-registered to the participant's high-resolution T1. Functional data underwent motion

412   correction (within- and between- scans), slice-time correction (pRF mapping runs only), and low-

413   pass filtering (60s). No spatial smoothing was done.

414   pRF model fitting: We fit population receptive field (pRF) model estimates following the event-related

415   paradigm introduced in Kay et al.[16]. This differs from many pRF implementations that fit the time

416   course of responses to sweeping bars or rings/wedges (e.g. [18,53], and instead allows us to combine

417   stimuli from different conditions (upright/inverted faces) in the same mapping run. mrVista functions

418   (https://github.com/vistalab/vistasoft) were used to estimate GLM betas for each voxel in each of 50

419   conditions (25 locations x upright/inverted) using the SPM (http://www.fil.ion.ucl.ac.uk/spm)

420   difference-of-gammas hemodynamic response function (HRF). Subsequent pRF modeling was done

421   using these beta amplitudes.

422   Population receptive field estimates were made independently for each voxel in each mapping

423   condition (upright/inverted) using the compressive spatial summation (CSS) model[29] that assumes a

424   circular 2D Gaussian pRF. The CSS model introduces a compressive power-law non-linearity
425   (exponent $n < 1$) that increases position invariance within the pRF and better models responses in
426   high-level visual areas[29] including face-selective regions[16]. We estimated pRFs independently in each
427   voxel and for each mapping condition by fitting six model parameters: gain, X, Y, σ, n. pRF size is
428   defined as $\sigma/\sqrt{n}$, and visualized in coverage density maps with a contour at radius = 2 x size from
429   the pRF center. Model fitting was performed using nonlinear optimization (MATLAB Optimization
430   Toolbox).

431   For quantification of pRF properties, we selected voxels with a goodness-of-fit $R^2 > 0.2$ in both
432   the upright and inverted conditions; in pooled-voxel visualizations (**Figure 3**), a threshold of $R^2 > 0.5$
433   was used for better visual clarity. Additionally, we trimmed voxels whose pRF fits were outside the
434   range of the mapping stimulus (central 11°) or within the fixation (central 0.4°), as well as those whose
435   size was estimated to be $< 0.1°$, as this typically corresponded to voxels that did not respond well
436   to our stimuli (e.g. uniformly low responses across the visual field).

437   In the main experiment, face position was coded using a binary mask determined by the
438   silhouette outline of the faces. These masks were drawn by hand for each face stimulus, and the
439   model fit was based on an average of the specific faces that the individual participant saw at each
440   of the 25 positions during his or her experimental session.

441   <u>Rescale + noise simulation (pRF experiment):</u> To elucidate how differential signal strength or model
442   goodness-of-fit may have contributed to the observed differences in inverted- and upright-mapped
443   pRF size and position estimates, we ran an iterative simulation on data from right hemisphere mFus-
444   faces, pooled across participants. As shown in **Figure S2**, simulated data for each voxel was
445   generated by taking beta estimates from the upright condition and matching beta range ('rescaling'
446   step) and model goodness-of-fit ('noise' step) to that of the inverted condition. The rescaling step
447   involved divisively scaling the range of the upright betas to match that of the inverted betas. The
448   noise step iteratively added Gaussian noise to the betas until the simulated goodness-of-fit R^2 was
449   within .01 of the R^2 in the inverted condition. To summarize the results of this simulation and compare
450   between upright, inverted, and simulated-data estimates, we used bootstrapping (1000 draws) to
451   estimate the median and 68% confidence interval of each model parameter of interest (Position Y,
452   size, and gain).

453   <u>Quantifying pRF size vs eccentricity:</u> pRF size linearly increases with eccentricity throughout the visual
454   hierarchy[15,16,59]. Thus, we compared the size vs. eccentricity relationship for upright and inverted

455   faces in each ROI. We performed this analysis in two ways: (i) pooling well-fitted voxels ($R^2 > 0.5$)

456   across participants in each ROI (**Figure 3**), and (ii) separately estimating the relationship for each

457   participant and ROI (**Figure S3**). The former produces more robust line fits, while the latter provides

458   an estimate of between-participant variability. Fitting was done by minimizing the L1 norm of the

459   residuals, e.g., the sum of the absolute values of the residuals. This solution was chosen for its

460   robustness to outliers in the underlying data.

461   <u>Visual field coverage:</u> Visual field coverage density plots for each visual region in **Figures 5 and S4**

462   were generated first for each participant, and then averaged across participants. Using a custom

463   Matlab bootstrapping procedure, these density plots represent the proportion of pRFs in a region

464   that overlap with each point in the visual field. Overlap is determined using a binary circular pRF at

465   the estimated center, and with a radius of $2\sigma/\sqrt{n}$[29]. This metric does not account for the Gaussian

466   profile of individual pRFs, but allows for greater interpretability when combining data across pRFs

467   and participants.

468   For each participant and ROI, density plots were generated by taking 1000 bootstrap samples

469   of 80% of voxels with replacement. Each voxel was represented with a circular binary mask as

470   described above, and coverage was computed by calculating the mean density across voxels for

471   each bootstrap draw. The average of these bootstrapped images is taken as the density coverage

472   for each participant. Participant-wise metrics, like FWHM area and center-of-mass (**Figure 5b**), are

473   computed from these images. Averages across participant-wise images are taken as overall

474   coverage density summaries (**Figure 5, Figure S6**); no rescaling or normalization is done, such that

475   plotting colors retain meaningful quantitative information about pRF coverage across the visual field.

476   <u>Maximal-overlap location simulation for behavioral experiment:</u> To evaluate the hypothesis that

477   inverted-face recognition would be improved at the retinotopic location that produced maximal

478   overlap with inverted-mapped pRFs, we ran a simulation using preliminary data from the first six

479   participants in the main experiment (**Figure 6b, Figure S7**). We pooled voxels from bilateral mFus-

480   faces across participants, and then took 1000 random samples of 100 voxels to generate a

481   simulated pRF coverage. We then took the dot product of this coverage with an averaged absolute-

482   contrast image of the internal features of the mapping faces at 3.2° diameter (such that the full face

483   image, not the internal features, was 3.2°), positioned at 0.1° intervals spanning the central 3° x 3°

484   of the visual field. For each simulated coverage map, this yielded an overlap metric at each of these

485   0.1° intervals (**Figure 6b, right panel**), and the maximal value was chosen. This yielded an average

486   maximal overlap when the face was positioned at 0.67° to the left and 0.79° below the center; the

487   leftward shift is a consequence of the fact that the right mFus-faces is larger than the left mFus-faces

488   in most participants, so that there are more voxels that have pRFs covering the left than right visual

489   field (see also[33]). For comparison, the same simulation using the internal features of an upright face

490   and the upright-mapped pRFs yielded a maximal-overlap position at 0.48° to the left and 0.58°

491   above the center. See also **Figure S7**, in which the maximal-overlap location is computed for each

492   of the 12 subjects in the full experiment.

493   Behavioral experiment: Following the simulation described above, we sought to probe recognition

494   performance for upright and inverted faces at three positions: the *lower left* location determined to

495   yield maximal overlap between inverted face features and inverted-mapped mFus-faces pRFs, a

496   mirrored *upper right* location that is equivalently far from the center, but predicted to have much

497   lower pRF overlap, and the *center* of the screen, where FIE is typically measured. The experiment

498   followed a standard behavioral face inversion paradigm[23], using a challenging recognition memory

499   task. In brief, during the behavioral task (**Figure 7a),** participants fixated on a small central bullseye

500   and saw a sequence of 3 faces each shown for 400ms, all upright or all inverted within each trial.

501   After an 800ms interval, a target face appeared at the same location. Participants were asked to

502   indicate whether the face was one of the faces in the prior triad (50% of trials) or not (50% of trials).

503   The experiments were run in the eye tracking lab, and participants' fixation was monitored with an

504   Eyelink 1000 as described below. All 12 of the scan participants took part in this experiment after

505   their scan session, and eye position was carefully monitored to ensure fixation. Data from 3

506   participants was excluded from the final analysis because they failed to maintain fixation on more

507   than 25% of all trials. Performance was measured for each condition (location x face orientation) as

508   a sensitivity index (d').

509   *Experimental design:* Prior to starting the behavioral experiment, participants received practice on

510   the task with performance feedback. Following the practice period, each participant completed 18

511   blocks of 20 trials. Position varied across blocks, such that each position (lower left, center, upper

512   right) was probed on 6 blocks. Within each block, half of the trials used upright faces and half used

513   inverted faces; trial ordering (i.e., inversion) was randomized. In total, each condition (face orientation

514   x position) was probed in 60 trials, and the experiment took ~30 minutes to complete.

515   *Trial structure:* Each trial (**Figure 7**) consisted of 3 faces presented for 400ms, followed by an

516   interstimulus interval of 800ms, and a 400ms target face. Participants were asked to respond via

517   keyboard whether the target face, which was marked with a thin green outline, had appeared in the

518   previous triad. Participants were given 2400ms from the appearance of the target face to respond.

519   Targets appeared on half of trials, randomly as the first, second, or third face in the triad. On trials

520   when the target was absent, it was replaced by a distractor face that was matched in general

521   appearance (e.g., 'blonde female with long hair') to the target, as maximally allowed by our set of 95

522   face identities. Face identities were the same as in the scan experiment. In the behavioral experiment,

523   we used front-view faces and also ± 15° viewpoint angle images of the same individuals to increase

524   the difficulty of the task. Trials were controlled so that the target image and its incidence in the triad

525   were always of different viewpoints. That is, an identical image was never used in the triad and as

526   the target. Each trial was preceded by a 2s fixation intertrial interval (ITI) during which eye position

527   was recorded to be used toward drift correction.

528   *Eyetracking:* Eyetracking during the behavioral experiment was performed using an Eyelink 1000

529   with a sampling rate of 1000Hz, and analyzed using a custom Matlab pipeline. Prior to starting the

530   experiment, participants received practice on the task with both performance feedback and real-

531   time eyetracking feedback: we plotted their current eye position as a series of red dots while they

532   performed the practice task at each of the probe positions. Participants were required to achieve

533   good fixation (< 0.5° deviation across trials in the second half of the practice) before they could

534   proceed to the main task. Additionally, to improve the validity of our eyetracking metrics, we

535   performed drift correction during each ITI; participants were aware of this and were instructed to

536   keep steady fixation even between trials. This allowed us greater confidence in the absolute values

537   of the tracked eye position on our displays, which was contained in the central ~5° of the visual field

538   in total. Following the experiment, we removed from the analysis any trials on which participants

539   broke fixation, and excluded 3 participants for whom this happened on > 25% of all trials. For the

540   remaining 9 participants, an average of 9.51% (s.d. = 4.65%) of trials across conditions were

541   removed as a result of breaking fixation.

542

543   Acknowledgements

545

546   Author Contributions

547 SP, KK, and KGS designed the experiment. SP and DF collected and analyzed data. SP and KK

548 wrote the analysis pipeline. KGS oversaw the data analyses. SP, KK, DF, and KGS wrote the

549 manuscript.

550

551 Data and Code Availability

552   Upon publication, data and code will be publicly available at github.com/VPNL/invPRF.

553

554 Declaration of Interests

555 The authors declare no competing interests.

References:

1. Konen, C. S. & Kastner, S. Two hierarchically organized neural systems for object information in human visual cortex. *Nat Neurosci* **11,** 224–231 (2008).
2. van Essen, D. C. & Gallant, J. L. Neural mechanisms of form and motion processing in the primate visual system. *Neuron* **13,** 1–10 (1994).
3. Ungerleider, L. 'What' and 'where' in the human brain. *Current opinion in neurobiology* **4,** 157–165 (1994).
4. Hubel, D. H. & Wiesel, T. N. Receptive fields and functional architecture in two nonstriate visual areas (18 and 19) of the cat. *Journal of Neurophysiology* **28,** 229–289 (1965).
5. Tong, F., Nakayama, K., Vaughan, J. T. & Kanwisher, N. Binocular Rivalry and Visual Awareness in Human Extrastriate Cortex. *Neuron* **21,** 753–759 (1998).
6. Grill-Spector, K., Knouf, N. & Kanwisher, N. The fusiform face area subserves face perception, not generic within-category identification. *Nat Neurosci* **7,** 555–562 (2004).
7. Gross, C. G., Rocha-Miranda, C. E. & Bender, D. B. Visual properties of neurons in inferotemporal cortex of the Macaque. *Journal of Neurophysiology* **35,** 96–111 (1972).
8. Dicarlo, J. J. & Cox, D. D. Untangling invariant object recognition. *Trends in Cognitive Sciences* **11,** 333–341 (2007).
9. Levy, I., Hasson, U., Avidan, G., Hendler, T. & Malach, R. Center–periphery organization of human object areas. *Nat Neurosci* **4,** 533–539 (2001).
10. Schwarzlose, R. F., Swisher, J. D., Dang, S. & Kanwisher, N. The distribution of category and location information across object-selective regions in human visual cortex. *Proceedings of the National Academy of Sciences* **105,** 4447–4452 (2008).
11. Arcaro, M. J., McMains, S. A., Singer, B. D. & Kastner, S. Retinotopic Organization of Human Ventral Visual Cortex. *J Neuro* **29,** 10638–10652 (2009).
12. Carlson, T., Hogendoorn, H., Fonteijn, H. & Verstraten, F. A. J. Spatial coding and invariance in object-selective cortex. *Cortex* **47,** 14–22 (2011).
13. Yue, X., Cassidy, B. S., Devaney, K. J., Holt, D. J. & Tootell, R. B. H. Lower-level stimulus features strongly influence responses in the fusiform face area. *Cerebral Cortex* **21,** 35–47 (2011).
14. de Haas, B. *et al.* Perception and Processing of Faces in the Human Brain Is Tuned to Typical Feature Locations. *J Neurosci* **36,** 9289–9302 (2016).
15. Dumoulin, S. O. & Wandell, B. A. Population receptive field estimates in human visual cortex. *Neuroimage* **39,** 647–660 (2008).
16. Kay, K. N., Weiner, K. S. & Grill-Spector, K. Attention Reduces Spatial Uncertainty in Human Ventral Temporal Cortex. *Current Biology* **25,** 595–600 (2015).
17. Gomez, J., Natu, V., Jeska, B., Barnett, M. & Grill-Spector, K. Development differentially sculpts receptive fields across early and high-level human visual cortex. *Nat Comms* **9,** 788 (2018).
18. Finzi, D. *et al.* Differential spatial processing in ventral and lateral face-selective regions is scaffolded by structural connections. *biorxiv.org* doi:10.1101/2020.07.06.190371
19. Haxby, J. V., Hoffman, E. A. & Gobbini, M. I. The distributed human neural system for face perception. *Trends in Cognitive Sciences* **4,** 223–233 (2000).
20. Parvizi, J. *et al.* Electrical Stimulation of Human Fusiform Face-Selective Regions Distorts Face Perception. *J Neuro* **32,** 14915–14920 (2012).
21. Jonas, J. *et al.* Focal electrical intracerebral stimulation of a face-sensitive area causes transient prosopagnosia. *Neuroscience* **222,** 281–288 (2012).
22. Weiner, K. S. *et al.* The mid-fusiform sulcus: a landmark identifying both cytoarchitectonic and functional divisions of human ventral temporal cortex. *Neuroimage* **84,** 453–465 (2014).
23. Yin, R. K. Looking at upside-down faces. *Journal of Experimental Psychology* **81,** 141–145 (1969).
24. Rossion, B. & Gauthier, I. How Does the Brain Process Upright and Inverted Faces? *Behavioral and Cognitive Neuroscience Reviews* **1,** 63–75 (2016).
25. Yovel, G. & Kanwisher, N. The neural basis of the behavioral face-inversion effect. *Current Biology* **15,** 2256–2262 (2005).

26. Kanwisher, N., Tong, F. & Nakayama, K. The effect of face inversion on the human fusiform face area. *Cognition* **68,** B1–11 (1998).
27. Van Belle, G., De Graef, P., Verfaillie, K., Rossion, B. & Lefèvre, P. Face inversion impairs holistic perception: evidence from gaze-contingent stimulation. *Journal of Vision* **10,** 10–10 (2010).
28. Rossion, B. Distinguishing the cause and consequence of face inversion: the perceptual field hypothesis. *Acta Psychol (Amst)* **132,** 300–312 (2009).
29. Kay, K. N., Winawer, J., Mezer, A. & Wandell, B. A. Compressive spatial summation in human visual cortex. *Journal of Neurophysiology* **110,** 481–494 (2013).
30. Haxby, J. V. *et al.* The effect of face inversion on activity in human neural systems for face and object perception. *Neuron* **22,** 189–199 (1999).
31. Le, R., Witthoft, N., Ben-Shachar, M. & Wandell, B. The field of view available to the ventral occipito-temporal reading circuitry. *Journal of Vision* **17,** 6 (2017).
32. Gilbert, C. & Bakan, P. Visual asymmetry in perception of faces. *Neuropsychologia* **11,** 355–362 (1973).
33. Yovel, G., Tambini, A. & Brandman, T. The asymmetry of the fusiform face area is a stable individual characteristic that underlies the left-visual-field superiority for faces. **46,** 3061–3068 (2008).
34. Andrews, T. J. & Ewbank, M. P. Distinct representations for facial identity and changeable aspects of faces in the human temporal lobe. *Neuroimage* **23,** 905–913 (2004).
35. Allison, T., Puce, A. & McCarthy, G. Social perception from visual cues: role of the STS region. *Trends in Cognitive Sciences* **4,** 267–278 (2000).
36. Pitcher, D. *et al.* The Human Posterior Superior Temporal Sulcus Samples Visual Space Differently From Other Face-Selective Regions. *Cereb Cortex* **30,** 778–785 (2020).
37. Henriksson, L., Mur, M. & Kriegeskorte, N. Faciotopy-A face-feature map with face-like topology in the human occipital face area. *Cortex* **72,** 156–167 (2015).
38. Taubert, J., Apthorp, D., Aagten-Murphy, D. & Alais, D. The role of holistic processing in face perception: Evidence from the face inversion effect. *Vision Research* **51,** 1273–1278 (2011).
39. Young, A. W., Hellawell, D. & Hay, D. C. Configurational information in face perception. *Perception* **16,** 747–759 (1987).
40. Valentine, T. Upside-down faces: a review of the effect of inversion upon face recognition. *Br J Psychol* **79 (Pt 4),** 471–491 (1988).
41. Tanaka, J. W. & Sengco, J. A. Features and their configuration in face recognition. *Mem Cognit* **25,** 583–592 (1997).
42. Rhodes, G., Brake, S. & Atkinson, A. P. What's lost in inverted faces? *Cognition* **47,** 25–57 (1993).
43. Avidan, G., Tanzer, M. & Behrmann, M. Impaired holistic processing in congenital prosopagnosia. *Neuropsychologia* **49,** 2541–2552 (2011).
44. Duchaine, B. C., Yovel, G., Butterworth, E. J. & Nakayama, K. Prosopagnosia as an impairment to face-specific mechanisms: Elimination of the alternative hypotheses in a developmental case. *Cognitive Neuropsychology* **23,** 714–747 (2006).
45. Behrmann, M. & Avidan, G. Congenital prosopagnosia: face-blind from birth. *Trends in Cognitive Sciences* **9,** 180–187 (2005).
46. Kay, K. N. & Yeatman, J. D. Bottom-up and top-down computations in word- and face-selective cortex. *eLife Sciences* **6,** 284 (2017).
47. Arcaro, M. J., Ponce, C. & Livingstone, M. The neurons that mistook a hat for a face. *eLife Sciences* **9,** 692 (2020).
48. Nishino, H. *et al.* Spatial Receptive Fields for Odor Localization. *Curr Biol* **28,** 600–608.e3 (2018).
49. Rust, N. C., Schwartz, O., Movshon, J. A. & Simoncelli, E. P. Spatiotemporal Elements of Macaque V1 Receptive Fields. *Neuron* **46,** 945–956 (2005).
50. Moser, E. I. *et al.* Grid cells and cortical representation. *Nature Reviews Neuroscience* **15,** 466–481 (2014).
51. Mancini, F., Haggard, P., Iannetti, G. D., Longo, M. R. & Sereno, M. I. Fine-Grained Nociceptive Maps in Primary Somatosensory Cortex. *J Neuro* **32,** 17155–17162 (2012).

52. van der Heijden, K., Rauschecker, J. P., de Gelder, B. & Formisano, E. Cortical mechanisms of spatial hearing. *Nature Reviews Neuroscience* **20,** 609–623 (2019).

53. Poltoratski, S. & Tong, F. Resolving the Spatial Profile of Figure Enhancement in Human V1 through Population Receptive Field Modeling. *J Neurosci* **40,** 3292–3303 (2020).

54. Klein, B. P., Harvey, B. M. & Dumoulin, S. O. Attraction of Position Preference by Spatial Attention throughout Human Visual Cortex. *Neuron* **84,** 227–237 (2014).

55. Witthoft, N., Poltoratski, S., Nguyen, M., BioRxiv, G. G.2016. Reduced spatial integration in the ventral visual cortex underlies face recognition deficits in developmental prosopagnosia. *biorxiv.org* doi:10.1101/051102

56. de Haas, B., Schwarzkopf, D. S., Anderson, E. J. & Rees, G. Perceptual load affects spatial tuning of neuronal populations in human early visual cortex. *Current Biology* **24,** R66–R67 (2014).

57. Stigliani, A., Weiner, K. S. & Grill-Spector, K. Temporal Processing Capacity in High-Level Visual Cortex Is Domain Specific. *J Neuro* **35,** 12412–12424 (2015).

58. Witthoft, N. *et al.* Where Is Human V4? Predicting the Location of hV4 and VO1 from Cortical Folding. *Cerebral Cortex* **24,** 2401–2408 (2013).

59. Wandell, B. A. & Winawer, J. Computational neuroimaging and population receptive fields. *Trends in Cognitive Sciences* **19,** 349–357 (2015).

Supplementary Information
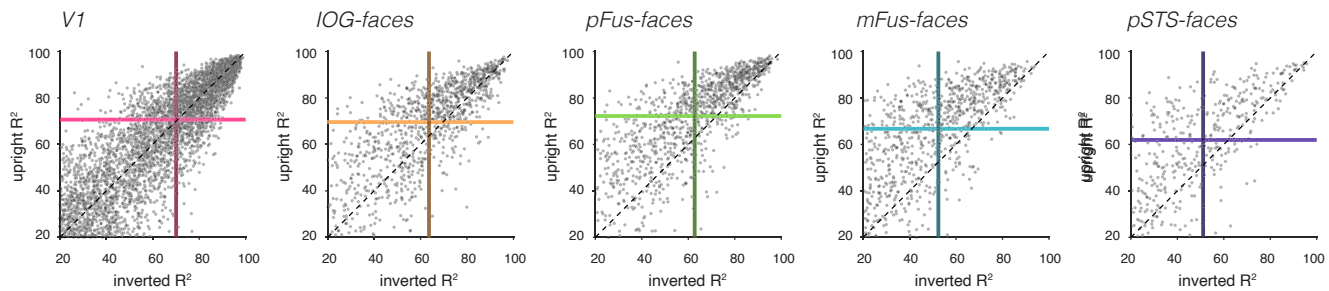
1 **Figure S1**. *Voxelwise comparison of model goodness-of-fit (R^2) across the visual hierarchy*. Voxels

2 are pooled across subjects, and the model goodness-of-fit for each voxel in the data set is shown.

3 Colored lines indicate the mean R^2 across voxels in the upright (horizontal, light colors) and inverted

4 (vertical, dark colors) mapping conditions. Model goodness-of-fit was equivalent across mapping

5 conditions in V1, but was lower in the inverted face condition in all face-selective areas (see also

6 Table S1).
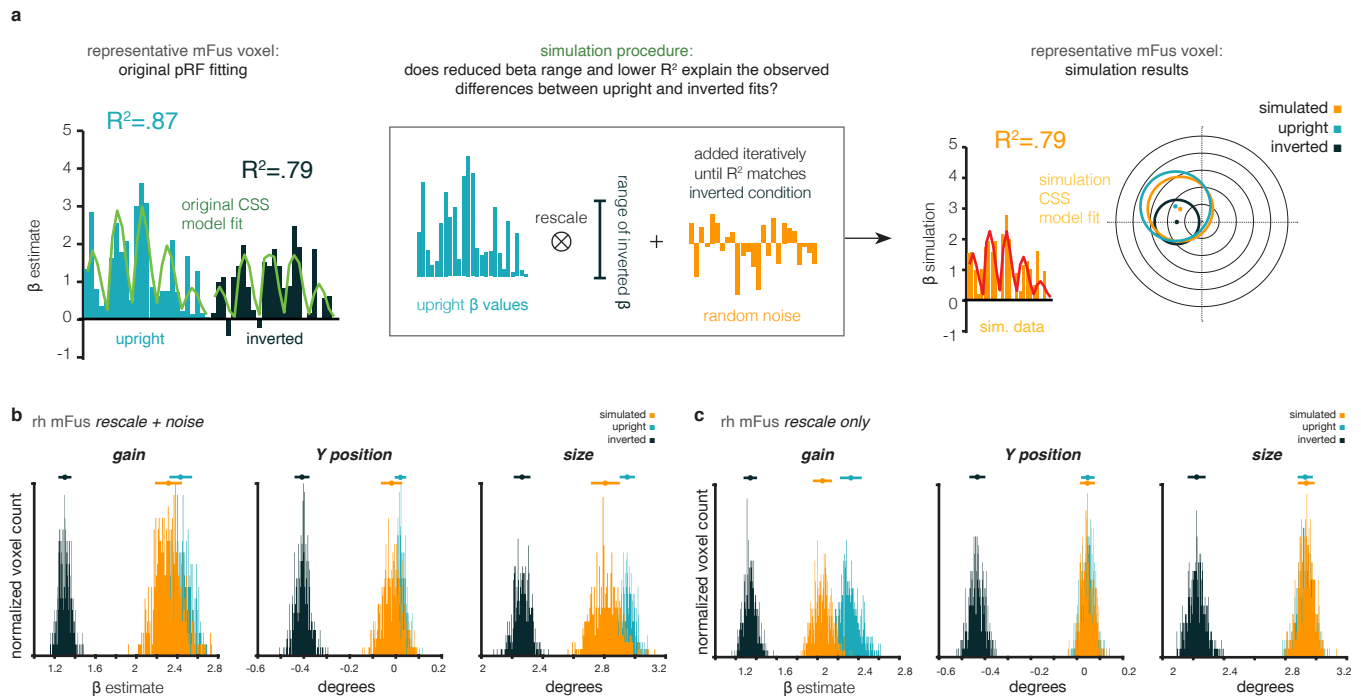
7   **Figure S2.** *Iterative noise simulation procedure and results*. **(a)** To what degree do the observed
8   differences in signal strength and R^2 between inverted and upright faces contribute to differences in
9   pRF position and size estimates? To determine this, we ran a simulation over mFus-faces voxels: for
10   each voxel, we took the 25 betas from the upright mapping condition (light blue), rescaled them to
11   match that the range of the 25 betas from the inverted mapping condition (dark blue; 'rescale' step),
12   and then iteratively added random noise ('noise' step) until the simulated model fit R^2 was equal to
13   the model R^2 in the inverted condition. In this example voxel, the simulated pRF fit (orange) much
14   more closely resembles the upright than the inverted condition fits. **(b)** Bootstrapped parameter
15   estimates for the rescaling + noise procedure described in **(a)**. While the simulated fits do show some
16   differences in size and gain compared to upright pRFs, the properties of the estimated pRFs much
17   more resembled the upright than inverted pRF and differences between upright and simulated pRFs
18   are substantially smaller in magnitude than what we observe for inverted faces, despite matching R^2
19   and signal level (βs). *Horizontal line markers* above distributions indicate median and 68% confidence
20   intervals around the estimates. **(c)** Results of an additional simulation which was the same as (b)
21   except that we omitted the noise step. Changing the response amplitude has minimal effect on the
22   size or position of pRFs. Together, these simulations allow us to conclude that differences in model
23   goodness-of-fit and weaker signal in response to inverted faces cannot fully account for the observed
24   differential pRF estimates for upright and inverted faces in the main experiment.
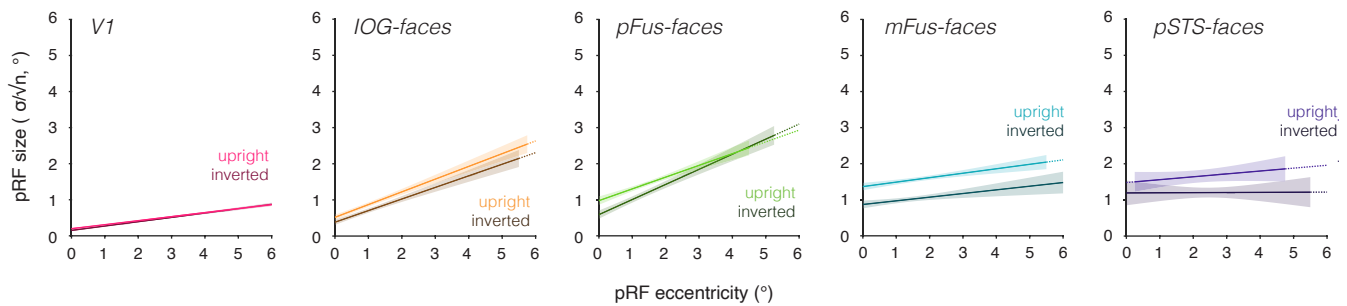
Figure S3. *Inverted-mapped pRFs are smaller in face-selective regions across eccentricity*. Related to **Figure 3**, here the size by eccentricity relationship in each ROI is estimated for each individual participant to provide an estimate of between-participant variability. Lines are generated by fitting least-squares lines to individual participant data and then averaging these fits. The line is subsampled at 0.25°, and shaded regions indicate ± SEM of the estimated model across participants, at eccentricities containing > 0 voxels in each individual participant. Decreased size for inverted-mapped pRFs is most apparent in mFus-faces, but is also evident in pFus- and pSTS-faces. We observed no reliable size differences across participants in V1 or IOG-faces. See also **Table S2**.

33 **Figure S4.** *Coverage metrics across participants.*

34 Related to **Figure 5**; we estimated for each ROI

35 and participant the area of coverage and the

36 center-of-mass of the coverage. Area (top) was

37 computed for the region at full-width-half-

38 maximum of the coverage plots shown in **Figure**

39 **5**. The center of mass is taken over the same

40 region, and has two parameters: a vertical Y

41 coordinate (middle) and horizontal X position

42 (bottom). *Dots:* individual participant metrics.

43 *Asterisks:* significant differences between upright

44 and inverted face conditions at p < .05, paired t-

45 test.
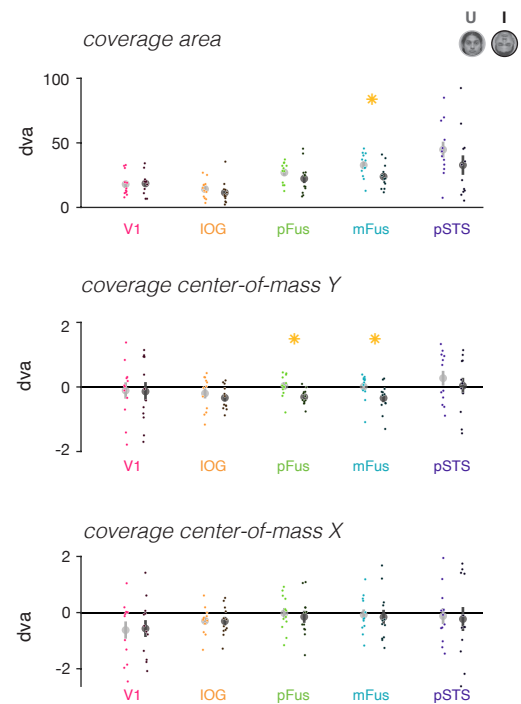
Figure S5. *Average pRF coverage density across participants by hemisphere*. Plotting conventions follow **Figure 5**; coverage density approximates the proportion of pRFs that overlap a given region in space. Mapping with upright faces (U) replicates previously reported properties of coverage in face-selective regions, including a strong contralateral representation of the fovea in pFus- and mFus-faces[1,2], and less foveal bias in pSTS-faces than in the ventral regions[3]. Mapping with inverted faces (I) yields significantly smaller full-width-half-max coverage area in lh-mFus, rh-mFus, lh-pSTS (t's(11) > 2.30, p's < 0.044) and a sign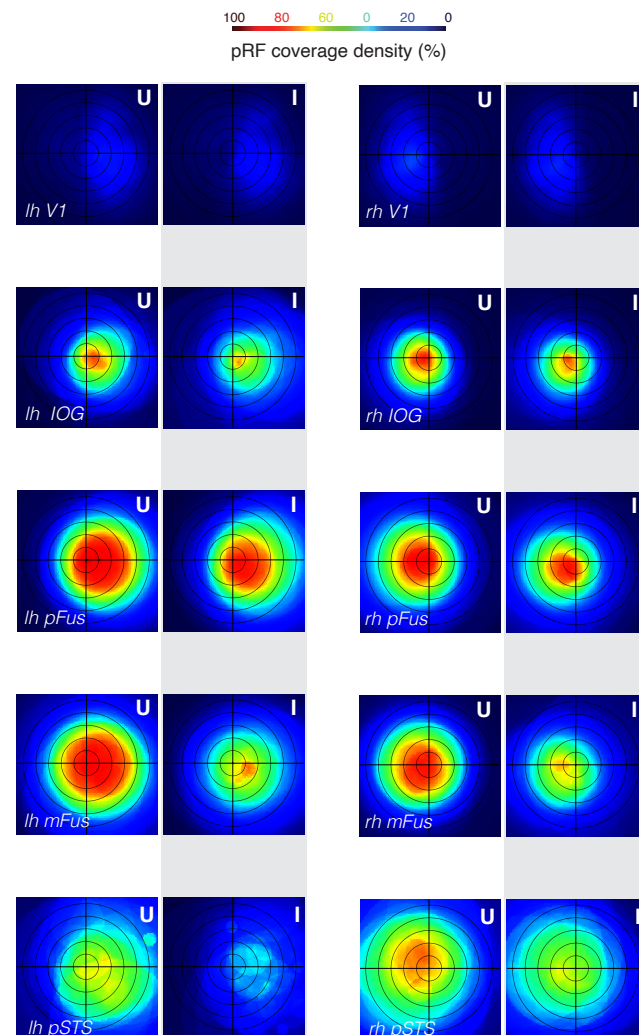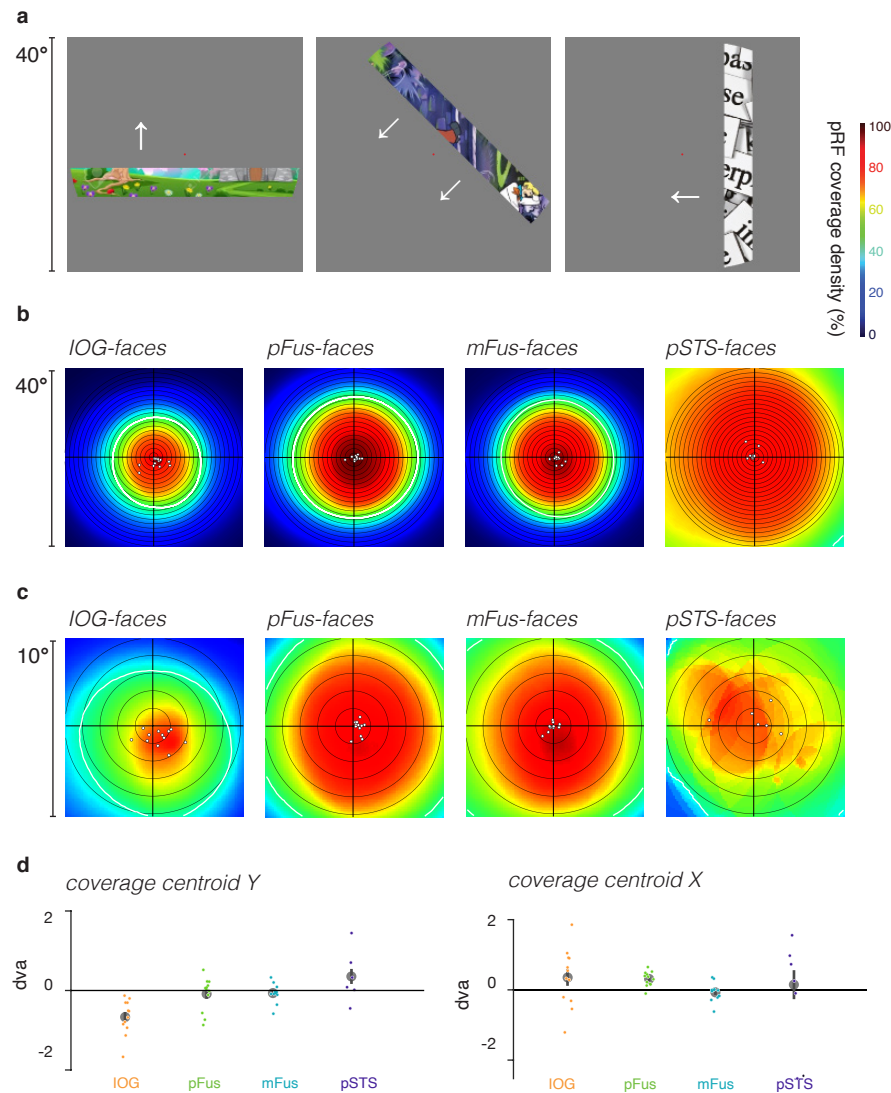ificant downward shift of the center-of-mass of the coverage in lh-pFus, rh-pFus, rh-IOG (t's(11) > 2.33, p's < 0.041). There were no significant differences in the horizontal location of the center-of-mass in any region and no significant changes to any of these parameters in either lh- or rh-V1.

Figure S6. *pRF mapping with cartoon stimuli yields pRF coverage consistent with upright-face mapping.* 12 participants in the current study took part in a separate pRF mapping experiment that used completely different mapping stimuli, mapping timing, and fitting procedures [3], allowing us to evaluate whether any non-optimal stimulus would yield a similar reported pattern of results as inverted faces, or whether changes in pRF properties and coverage are specific to face inversion. For brevity, here we refer to 'toonotopy data,' as the subset of data from Finzi et al.[3] for the participants in the current study (N = 12), analyzed using the same participant-wise ROIs as in the current study. Unlike the current study, for the toonotopy data we fit the CSS pRF model over mean BOLD timecourses (procedure: http://cvnlab.net/analyzePRF) rather than to estimated beta weights for discrete mapping stimuli. To better compare the widefield data to the current study, we analyzed both the full toonotopy datasets of the 12 subjects, and also a subset of the data matched to the

80   spatial extent of the mapping stimuli in the current study (eccentricity < 5°, size < 10°). **(a)** Toonotopy

81   mapping stimuli were comprised of rapidly-presented cartoon frames (8Hz presentation) within a bar

82   aperture that swept the visual field in 8 directions, and covered the central 40° of the visual field over

83   the course of an fMRI run. **(b)** For the 12 participants in the current study, the average pRF coverage

84   in face-selective ROIs from the full toonotopy data. As in Figure 5, *white contour:* the average FWHM

85   of coverage density. *White dots:* individual participants' FWHM center-of-mass position. As

86   expected, the widefield nature of the toonotopy stimuli of Finzi and colleagues [3] allow us to measure

87   pRFs beyond the ~10° extent of the mapping stimuli used in our main experiment. To draw better

88   comparison to the current study, **(c)** presents equivalent coverage plotting of the subset of subjects'

89   pRFs in the toonotopy data matched to the ~10° extent of the mapping stimuli of the current study.

90   Cross-participant metrics of this coverage in the face-selective regions are shown in **(d)**. While we

91   cannot quantitatively compare visual field coverage area between the two studies due to the differing

92   spatial extent of the mapping stimuli, we see qualitatively similar coverage in response to the

93   toonotopy mapping stimuli and the upright face mapping in the current study. That is, the visual field

94   coverage was foveally biased in ventral face-selective regions and progressively increased from IOG-

95   to pFus- to mFus-faces. We see no evidence of a downward shift in the center-of-mass for pFus-

96   and mFus-faces as observed with inversion.
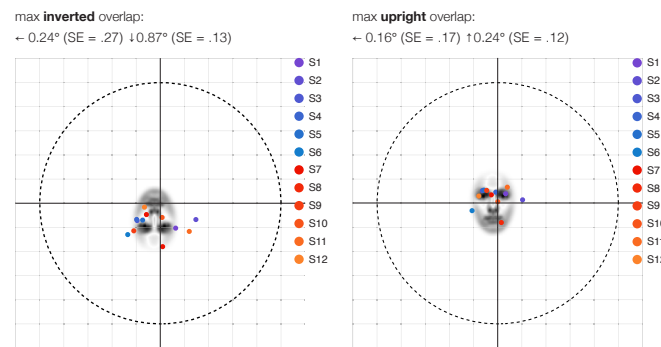
97   **Figure S7.** *Positions in the visual field yielding maximal pRF overlap for each of our participants.* As

98   described in **Figure 6**, we sought to determine the position in space at which the internal features of

99   an 3.2° upright face (left panel) maximally overlapped with upright-mapped pRFs, and the position

100   in space at which the internal features of an inverted face (right panel) maximally overlapped with

101   inverted-mapped pRFs. The bootstrapping simulation used to determine the positions at which faces

102   were shown in the behavioral experiment was based on pooled preliminary data from the first 6

103   participants in the experiment (*S1-6,* cool colors). We repeated the simulation procedure on an

104   individual-participant basis (1000 draws per participant, 80% of voxels sampled with replacement),

105   including the participants whose pRF data was collected after the initial simulation (S7-12, warm

106   colors). *Dots:* represent bootstrapped estimates for individual participants; *face images* in each panel

107   are shown at the mean maximal overlap position in each condition. Overall, while some inter-

108   participant variability is present, the mean positions determined across all participants align well with

109   the position used in the behavioral experiment, such that maximal overlap for inverted faces occurs

110   in the lower left visual field, while maximal overlap for upright faces occurs is more central and trends

111   slightly to the upper left.
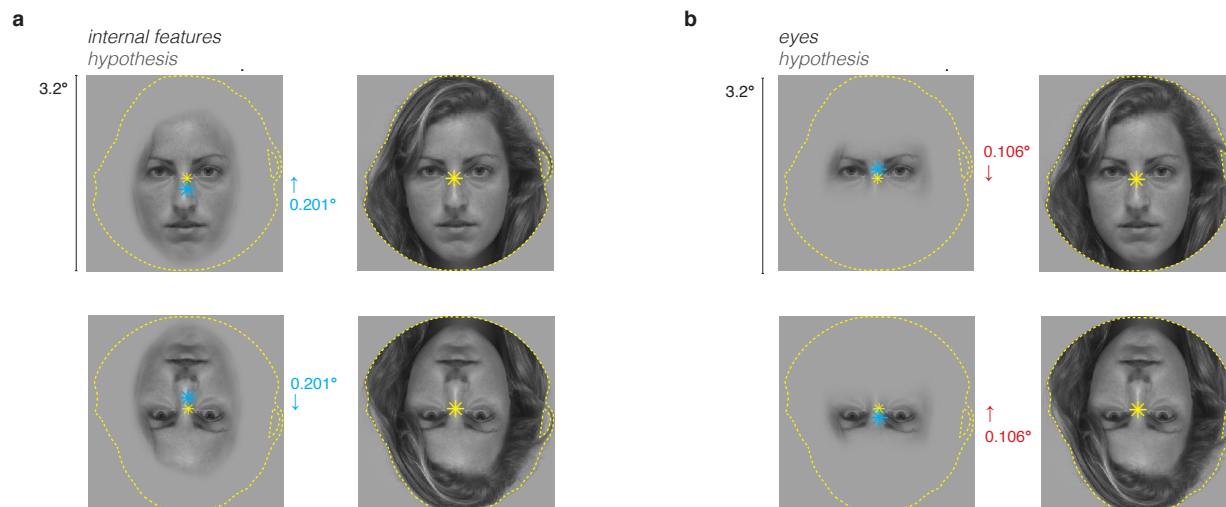
Figure S8. *The location of internal face features may predict the magnitude and directionality of the reported pRF shifts with face inversion.* In this schematic, we present two hypotheses that examine how the location of internal features (left) or eyes alone (right) may contribute to the positional shifts in pRFs due to face inversion. In the pRF model, each individual face location is coded as a binary silhouette of the full-face image that participants saw, here shown via *yellow dashed line*. Because pRF model fitting relies on computing the dot product of a hypothesized Gaussian pRF and this binary image, we can infer that the center-of-mass of the binary silhouette is an important factor in determining the estimated pRF position. The weighted centroid of the binary silhouette for the sample image is marked by *yellow asterisks* in all panels. To account for the observed shift in Y position in response to face inversion, we evaluated the (a) *Internal features hypothesis*: the position of the internal face features, rather than the full face image, drives spatial responses in face-selective regions. This can be approximated by calculating the weighted centroid of the internal features of this face for the upright and inverted conditions (left panel, *blue asterisks*). Correspondingly, if the internal features rather than the entire face determine the location of the pRF in both the upright and inverted mapping conditions (blue), our implemented pRF model (yellow) will yield position estimates ~0.2° above the location of the internal features of the face in the upright condition, and ~0.2° below in the inverted condition. The overall consequence of these offsets would be a ~0.4° lower pRF estimate in the inverted condition than in the upright condition, which is similar to the magnitude and directionality of effects we observed in pFus- and mFus-faces. (b) *Eyes hypothesis:* the position of the eyes alone, rather than the full face image, drives spatial responses in face-selective regions. This can be approximated by calculating the weighted centroid of the eyes of this face for the upright and

133   inverted conditions (right panel, *blue asterisks*). Correspondingly, if the location of the eyes within the

134   mapping stimulus drives spatial responses in a voxel, we would expect to see (i) smaller difference

135   between the upright- and inverted-mapped pRF estimates, as the centroid of the eye region of our

136   face images (blue asterisk) is typically well-aligned with the centroid of the binary silhouette (yellow),

137   and (ii) inverted-mapped pRFs that are shifted upward ~0.2° relative to the upright-mapped

138   condition. This is a poorer match to the effects of face inversion observed experimentally; neither

139   manipulation predicts the observed differences in pRF gain or size.

140 **Table S1.** Results of three-way repeated measures analyses of variance (ANOVA) with factors of

141 hemisphere (right/left), region of interest (IOG-/pFus-/mFus-/pSTS-faces), and condition

142 (upright/inverted) for each pRF mean parameter estimate across participants. One participant had

143 insufficient lh-pSTS data (0 voxels above threshold) and was excluded from this analysis, leaving N

144 = 11. Related to **Figure 2**.

center position Y estimate		
significant main effects:		
condition	F(1,10) = 8.82	p = 0.014
significant interactions:		
none		
not significant:		
hem	F(1,10) = 0.10	p = 0.758.
ROI	F(3,30) = 0.10	p = 0.681
hem x ROI	F(3,30) = 0.55	p = 0.649
hem x condition	F(1,10) = 2.61	p = 0.137
ROI x condition	F(3,30) = 0.83	p = 0.489
hem x ROI x condition	F(3,30) = 0.67	p = 0.577

size estimate (σ / √ n)		
significant main effects:		
ROI	F(3,30) = 7.13	p = 0.001
condition	F(1,10) = 56.07	p < 0.001
significant interactions:		
hem x ROI	F(3,30) = 3.36	p = 0.032
ROI x condition	F(3,30) = 7.53	p = 0.001
not significant:		
hem	F(1,10) = 0.01	p = 0.920.
hem x condition	F(1,10) = 1.18	p = 0.303
hem x ROI x condition	F(3,30) = 2.64	p = 0.068.

center position X estimate		
significant main effects:		
hem	F(1,10) = 375.60	p < 0.001
significant interactions:		
hem x ROI	F(3,30) = 4.49	p = 0.010.
not significant:		
ROI	F(3,30) = 2.15	p = 0.114
condition	F(1,10) = 0.31	p = 0.592
hem x condition	F(1,10) = 1.16	p = 0.306
ROI x condition	F(3,30) = 0.39	p = 0.763
hem x ROI x condition	F(3,30) = 0.06	p = 0.982

gain estimate		
significant main effects:		
none		
significant interactions:		
ROI x condition	F(3,30) = 4.66	p = 0.009
not significant:		
hem	F(1,10) = 2.77	p = 0.127
ROI	F(3,30) = 0.61	p = 0.615
condition	F(1,10) = 2.07,	p = 0.181
hem x ROI	F(3,30) = 0.75	p = 0.530
hem x condition	F(1,10) = 0.05	p = 0.824
hem x ROI x condition	F(3,30) = 0.36	p = 0.781

model R^2		
significant main effects:		
ROI	$F(3,30) = 17.59$	$p < 0.001$
condition	$F(1,10) = 80.36$	$p < 0.001$
significant interactions:		
ROI x condition	$F(3,30) = 7.47$	$p = 0.001$
not significant:		
hem	$F(1,10) = 0.19$	$p = 0.669$
hem x ROI	$F(3,30) = 1.13$	$p = 0.352$
hem x condition	$F(1,10) = 1.28$	$p = 0.283.$
hem x ROI x condition	$F(3,30) = 0.23$	$p = 0.872$

145   **Table S2.** Results of two-way repeated measures ANOVA with factors of ROI (lOG-/pFus-/mFus-

146   /pSTS-faces) and condition (upright/inverted) on the relationship of size and eccentricity in bilateral

147   face-selective areas (related to **Figure S3**) across participants (N=12).

fit line slope		
significant main effects:		
ROI	F(3) = 3.23	p = 0.035
significant interactions:		
none		
not significant:		
condition	F(1) = 0.04	p = 0.853
ROI x condition	F(3) = 0.45	p = 0.719

fit line intercept		
significant main effects:		
condition	F(1) = 9.08	p = 0.012
significant interactions:		
none		
not significant:		
ROI	F(3) = 2.80	p = 0.055
ROI x condition	F(3) = 0.58	p = 0.633

148   **Table S3.** Results of two-way repeated measures ANOVA on the individual participants' (N=12)
149   coverage density metrics in bilateral face-selective areas (related to **Figure 5**).

coverage density FWHM area		
significant main effects:		
ROI	$F(3) = 3.10$	$p = 0.040$
condition	$F(1) = 7.21$	$p = 0.021$
significant interactions:		
none		
not significant:		
ROI x condition	$F(3) = 0.39$	$p = 0.758$

center of mass Y position		
significant main effects:		
ROI	$F(3) = 3.10$	$p = 0.040$
condition	$F(1) = 7.21$	$p = 0.021$
significant interactions:		
none		
not significant:		
ROI x condition	$F(3) = 0.39$	$p = 0.758$

center of mass X position		
significant main effects:		
none		
significant interactions:		
none		
not significant:		
ROI	$F(3) = 0.18$	$p = 0.907$
condition	$F(1) = 0.25$	$p = 0.627$
ROI x condition	$F(3) = 0.05$	$p = 0.984$

Supplementary References

1. Kay, K. N., Weiner, K. S. & Grill-Spector, K. Attention Reduces Spatial Uncertainty in Human Ventral Temporal Cortex. Current Biology 25, 595–600 (2015).
2. Gomez, J., Natu, V., Jeska, B., Barnett, M. & Grill-Spector, K. Development differentially sculpts receptive fields across early and high-level human visual cortex. Nat Comms 9, 788 (2018).
3. Finzi, D. et al. Differential spatial processing in ventral and lateral face-selective regions is scaffolded by structural connections. biorxiv.org doi:10.1101/2020.07.06.190371