Morphing between expressions dissociates continuous from categorical representations of facial expression in the human brain

Richard J. Harris, Andrew W. Young, and Timothy J. Andrews[1]

Department of Psychology and York Neuroimaging Centre, University of York, York YO10 5DD, United Kingdom

Whether the brain represents facial expressions as perceptual continua or as emotion categories remains controversial. Here, we measured the neural response to morphed images to directly address how facial expressions of emotion are represented in the brain. We found that face-selective regions in the posterior superior temporal sulcus and the amygdala responded selectively to changes in facial expression, independent of changes in identity. We then asked whether the responses in these regions reflected categorical or continuous neural representations of facial expression. Participants viewed images from continua generated by morphing between faces posing different expressions such that the expression could be the same, could involve a physical change but convey the same emotion, or could differ by the same physical amount but be perceived as two different emotions. We found that the posterior superior temporal sulcus was equally sensitive to all changes in facial expression, consistent with a continuous representation. In contrast, the amygdala was only sensitive to changes in expression that altered the perceived emotion, demonstrating a more categorical representation. These results offer a resolution to the controversy about how facial expression is processed in the brain by showing that both continuous and categorical representations underlie our ability to extract this important social cue.

face | temporal lobe | fMRI

The ability to visually encode changes in facial musculature that reflect emotional state is essential for effective social communication (1). Models of face processing have proposed either that the perception of facial expression is dependent on a continuous neural representation of gradations in expression along critical dimensions (2, 3) or that it is based primarily on the assignment of expressions to discrete emotion categories (4, 5).

Although they are usually treated as incompatible opposites, there is evidence consistent with both the continuous and categorical approaches to facial expression perception (6). The categorical perspective is based on the idea that distinct neural or cognitive states underpin a set of basic facial expressions (5). Evidence for categorical perception of expression is shown by the consistency with which the basic emotions are recognized (7) and the greater sensitivity to changes in facial expression that alter the perceived emotion (8, 9). In contrast, continuous or dimensional models are better able to explain the systematic confusions that occur when labeling facial expressions (2) and can account for variation in the way that basic emotions are expressed (10) and the fact that we are readily able to perceive differences in intensity of a given emotional expression (11, 12).

The aim of this study is to ask whether primarily categorical or continuous representations of facial expression are used in different regions of the human brain. Neuroimaging studies have identified a number of face-selective regions that are involved in the perception of facial expression (13–16). The occipital face area (OFA) is thought to be involved in the early perception of facial features and has a postulated projection to the posterior superior temporal sulcus (pSTS). The connection between the OFA and pSTS is thought to be important in processing dynamic changes in the face, such as changes in expression and gaze,

which are important for social interactions (17–21). Information from the pSTS is then relayed to regions of an extended face-processing network, including the amygdala, for further analysis of facial expression.

To determine how different regions in the face-processing network are involved in the perception of emotion, we compared the responses to faces that varied in facial expression. In experiment 1, we compared the response to faces that changed in both expression and identity. Our aim was to determine which regions responded selectively to changes in facial expression—i.e., those that responded more strongly to changes in expression than to changes in identity. In experiment 2, we used morphs between different images of facial expressions to ask whether the neural response in those face regions reflected a continuous or categorical representation of emotion. To achieve this goal, the face images we used could be physically identical ("same expression"), could differ in physical properties without crossing the category boundary ("within-expression change"), or could differ in physical properties and cross the category boundary ("between-expression change"). Importantly, both the within-and between-expression conditions involved an equivalent 33% shift along the morphed continuum. Brain regions that hold a more categorical perception of expression should be sensitive to between-expression changes in expression, but not within-expression changes. However, regions with a continuous representation should be equally sensitive to both between- and within-expression changes.

Results

Experiment 1. The aim of experiment 1 was to determine which regions in the face-selective network were sensitive to changes in facial expression. There were four conditions: (*i*) same expression, same identity; (*ii*) same expression, different identity; (*iii*) different expression, same identity; and (*iv*) different expression, different identity (Fig. 1*A*). The peak responses of the face-selective regions, defined in an independent localizer scan (Fig. S1 and Table S1), were analyzed by using a $4 \times 2 \times 2$ ANOVA with region [pSTS, amygdala, fusiform face area (FFA), OFA], expression (same, different), and identity (same, different) as the factors. There were significant effects of expression [$F_{(1, 13)} = 4.46$, $P = 0.05$] and region [$F_{(3, 39)} = 48.26$, $P < 0.0001$], but not identity [$F_{(1, 13)} = 2.52$, $P = 1.14$]. There was also a significant interaction between region × expression [$F_{(3, 39)} = 12.73$, $P < 0.0001$]. Therefore, to investigate which face-selective regions were sensitive to expression, we next considered the response in each individual region of interest (ROI).
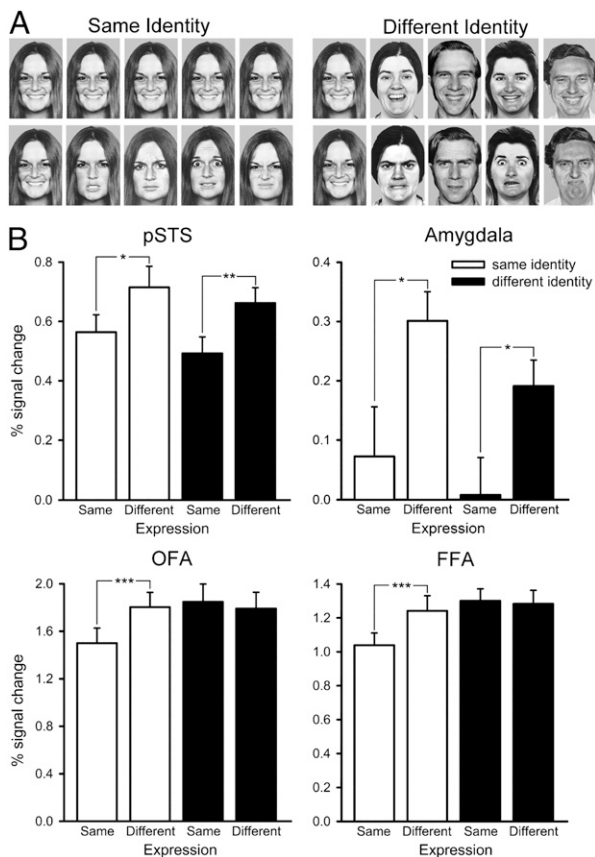
Fig. 1. Experiment 1: Sensitivity to changes in expression and identity. (*A*) Images from the four conditions. (*Upper*) Same expression, same identity (*Left*); same expression, different identity (*Right*). (*Lower*) Different expression, same identity (*Left*); different expression, different identity (*Right*). (*B*) Peak responses to the different conditions in the pSTS (*Upper Left*), amygdala (*Upper Right*), OFA (*Lower Left*), and FFA (*Lower Right*). The graphs show that face-selective regions in the pSTS and the amygdala responded selectively to changes in facial expression, independent of changes in identity. Error bars represent SE. *$P < 0.05$; **$P < 0.01$; ***$P < 0.001$. Images of faces are from ref. 47.

Fig. 1*B* shows the response from the pSTS in experiment 1. A 2×2 ANOVA with the factors expression (same, different) and identity (same, different) revealed a significant effect of expression [$F_{(1, 17)} = 12.84$, $P = 0.002$], but not identity [$F_{(1, 17)} = 1.98$, $P = 0.18$]. There was no significant interaction between expression and identity [$F_{(1, 17)} = 0.04$]. The effect of expression was driven by a significantly larger response to the different-expression conditions compared with the same-expression conditions in both the same-identity [$t_{(17)} = 2.20$, $P = 0.04$] and different-identity [$t_{(17)} = 2.75$, $P = 0.01$] conditions.

The amygdala revealed a similar pattern of results to that found in the pSTS (Fig. 1*B*). A 2×2 repeated-measures ANOVA found a significant main effect of expression [$F_{(1, 15)} = 11.13$, $P = 0.01$] but not identity [$F_{(1, 15)} = 0.09$]. There was no significant interaction between expression and identity [$F_{(1, 15)} = 2.69$, $P = 0.12$]. Again, the main effect of expression was driven by the significantly larger response to different expression compared with same expression in the same-identity [$t_{(14)} = 2.18$, $P = 0.05$] and the different-identity [$t_{(14)} = 2.23$, $P = 0.04$] conditions.

The FFA was sensitive to both changes in expression and identity. A 2×2 ANOVA revealed significant main effects of expression [$F_{(1, 19)} = 18.06$, $P < 0.0001$] and identity [$F_{(1, 19)} = 4.53$, $P = 0.05$]. There was also a significant interaction between expression and identity [$F_{(1, 19)} = 7.18$, $P = 0.02$]. The main effect of expression was due to a bigger response to the different-

expression condition compared with the same-expression condition for same-identity faces [$t_{(19)} = 4.39$, $P < 0.0001$]. However, there was no significant difference between the different- and same-expression conditions for different-identity faces [$t_{(19)} = 0.86$].

The OFA showed a similar pattern of response to that found in FFA. There were significant main effects of expression [$F_{(1, 19)} = 12.71$, $P < 0.002$] and identity [$F_{(1, 19)} = 9.91$, $P = 0.01$]. There was also a significant interaction between expression and identity [$F_{(1, 19)} = 8.58$, $P = 0.01$]. We found a significantly larger response to different-expression condition compared with same-expression condition for the same-identity [$t_{(19)} = 4.24$, $P < 0.0001$], but not for the different-identity conditions [$t_{(19)} = 0.16$].

The results from experiment 1 therefore show selectivity to changes in facial expression (stronger responses to changes in expression than to changes in identity) for pSTS and amygdala. However, because our aim was to investigate Haxby and colleagues' neural model of face perception (15), our ROI analysis was restricted to face-selective regions defined by the independent localizer. Thus, a group analysis was run to determine whether there were other regions that were selective for differences in facial expression, but were not face selective. Consistent with the ROI analysis, we found that the pSTS, amygdala, FFA, and OFA showed increased responses to different expression conditions compared with the same expression conditions (Fig. S2). However, we did not find regions outside these core face-selective regions that showed selectivity for expression (Table S2).

Experiment 2. Experiment 2 used morphed images to determine whether different face regions have a continuous and categorical representation of emotion. Continua for the experiment were generated by morphing between different expression images (Fig. 2). Validation of the morphing procedure was demonstrated in two behavioral experiments. First, we performed an expression-classification experiment. Four images were selected along the
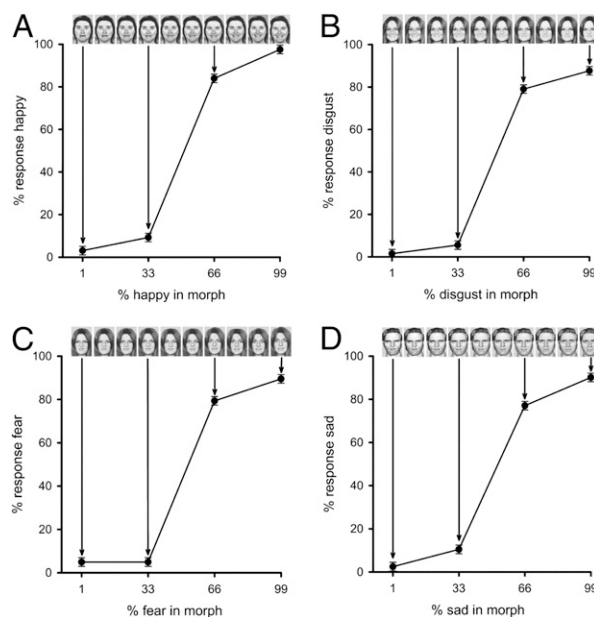
Fig. 2. Experiment 2: Behavioral results from expression-classification experiment. Performance was determined for four morph continua: fear to happy (*A*), happy to disgust (*B*), disgust to fear (*C*), and disgust to sad (*D*). The *x* axis shows the four morph levels from the continua that were used in the experiment. The graph represents the proportion of participant's responses that used the name given on the *y* axis, averaged across participants. A clear discontinuity is evident in the perception of emotion near the midpoint of the morphed continuum, which provides support for a categorical representation of facial expression. Images of faces are from ref. 47.

appropriate morph continua (99%, 66%, 33%, and 1%), and participants were asked to make a five-alternative forced choice (5AFC) for the following expression continua: happy–fear, disgust–happy, sad–disgust, fear–disgust. Fig. 2 shows that for each set of images, there was a clear discontinuity in the perception of emotion near the midpoint of the morphed continuum.

The second behavioral experiment involved a more stringent test of categorical perception using a same–different task. Participants were presented with two sequentially presented faces with the same identity and had to judge whether the images were identical or different. There were three conditions as follows: (i) same emotion (99% and 99% or 66% and 66%), (ii) within emotion (99% and 66% images), and (iii) between emotion (66% and 33% images). We found that the between-emotion condition was judged as different more often relative to the within-emotion condition [$t_{(12)} = 6.47$, $P = <0.001$; Fig. S3]. Moreover, participants responded faster on correct responses to the between-emotion compared with the within-emotion condition [$t_{(2,24)} = 4.19$, $P = <0.001$]. These results show that facial expressions that differ in perceived emotion are discriminated more easily than facial expressions that are perceived to convey the same emotion. This finding is widely considered to form the strongest test for behavioral evidence of categorical perception (12).

For the functional MRI (fMRI) experiment, stimuli in a block were selected from three faces along the morphed continuum (99%, 66%, and 33%). The within-emotion condition used two faces from the morph continua that were on the same side of the category boundary (99% and 66%). The between-emotion condition used two faces along the morph continua that crossed the category boundary (66% and 33%). Importantly, the physical difference between images was therefore matched across within- and between-emotion conditions (both 33% difference). The same emotion condition had two identical images (99% and 99% or 66% and 66%). The identity of the faces was also varied to give six conditions as follows: (i) same expression, same identity; (ii) within expression, same identity; (iii) between expression, same identity; (iv) same expression, different identity; (v) within expression, different identity; and (vi) between expression, different identity (Fig. 3). The peak responses of the face-selective regions were analyzed by using a 4 × 3 × 2 ANOVA with region

(pSTS, amygdala, FFA, or OFA), expression (same, within, or between), and identity (same or different) as the factors. There were significant effects of expression [$F_{(2, 24)} = 15.39$, $P < 0.0001$] and region [$F_{(3, 36)} = 49.12$, $P < 0.0001$], but not identity [$F_{(1, 12)} = 3.88$, $P = 0.07$]. There was also a significant interaction between region × expression [$F_{(6, 72)} = 2.43$, $P = 0.03$]. Therefore, to investigate which face-selective regions were sensitive to expression—and in what way each was sensitive to differences in expression—we next considered the response in each individual ROI.

Fig. 3C shows the response from the pSTS in experiment 2. To determine the sensitivity of the pSTS to these changes, we conducted a 3 × 2 ANOVA with the factors expression (same, within, or between) and identity (same or different). This analysis revealed a significant effect of expression [$F_{(2, 32)} = 13.19$, $P < 0.0001$] but no significant effect of identity [$F_{(1, 16)} = 0.60$] or a significant interaction [$F_{(2, 32)} = 0.25$]. The main effect of expression for the same-identity conditions was due to significantly bigger responses to the within-expression [$t_{(16)} = 3.00$, $P = 0.01$] and between-expression [$t_{(16)} = 3.88$, $P = 0.001$] conditions compared with the same-expression condition. There was no significant difference between the within- and between-expression conditions [$t_{(16)} = 0.40$]. A similar pattern of results was seen for the different-identity conditions. There were significantly larger responses to the within-expression [$t_{(16)} = 2.49$, $P = 0.02$] and between-expression [$t_{(16)} = 3.17$, $P = 0.01$] conditions compared with the same-expression condition. However, there was no significant difference in the response to the within- and between-expression conditions [$t_{(16)} = 0.22$]. This equivalent sensitivity to both within- and between-expression changes suggests that the pSTS has a continuous representation of expression.

In contrast to the pSTS, the amygdala was only sensitive to between-emotion changes in expression. An ANOVA revealed a significant main effect of expression [$F_{(2, 48)} = 22.52$, $P < 0.0001$] but not identity [$F_{(1, 24)} = 4.03$, $P = 0.06$], and there was also no significant interaction [$F_{(2, 48)} = 1.49$, $P = 0.29$]. For the same-identity conditions, there was no significant difference between the same- and within-expression conditions [$t_{(20)} = 1.61$, $P = 0.12$]. However, there was a significant difference between the same- and between-expression conditions [$t_{(20)} = 4.86$, $P < 0.0001$]

Fig. 3. Experiment 2: Continuous and categorical representations of expression. (A) Images from the three expression conditions with the same identity. (B) Images from the three expression conditions with different identities. (C) Peak responses to the different conditions in the pSTS and amygdala. The results show that the pSTS was equally sensitive to all changes in facial expression, consistent with a continuous representation. In contrast, the amygdala was only sensitive to changes in expression that altered the perceived emotion, demonstrating a more categorical representation of emotion. *$P < 0.05$, **$P < 0.01$, ***$P < 0.001$. Images of faces are from ref. 47.

and between the within- and between-expression conditions [$t_{(20)}$ = 4.62, $P < 0.0001$]. There was a similar pattern for the different-identity conditions. There was no significant difference between the same- and within-expression conditions [$t_{(20)} = 0.84$], but there was a larger response to the between-expression condition compared with the within-expression conditions [$t_{(20)} = 3.58, P < 0.001$]. There was also a larger response to the between- and within-expression conditions [$t_{(20)} = 2.06, P = 0.05$]. These results show that the amygdala is more sensitive to changes in expression that cross an emotion category boundary.

In the FFA (Fig. S4), there was significant main effects of expression [$F_{(2, 48)} = 6.36, P = 0.004$] and identity [$F_{(1, 24)} = 9.29, P = 0.01$], but there was no significant interaction between expression and identity [$F_{(2, 48)} = 2.54, P = 0.09$]. For the same-identity conditions, there was no significant difference between the same- and within-expression conditions [$t_{(24)} = 1.71, P = 0.10$]. There was also no significant difference between the within- and between-expression conditions [$t_{(24)} = 1.38, P = 0.06$]. However, there was a significant difference between the same- and between-expression conditions [$t_{(24)} = 2.96, P = 0.01$]. For the different-identity conditions, there were no significant difference between the same-expression and either the within-expression [$t_{(24)} = 1.88, P = 0.07$] or between-expression [$t_{(24)} = 0.65$] conditions. There was also no difference in response between the between- and within-expression conditions [$t_{(24)} = 1.50, P = 0.15$].

In the OFA, there was a significant main effect of expression [$F_{(2, 48)} = 8.53, P = 0.001$] and identity [$F_{(1, 24)} = 7.77, P = 0.01$]. There was also a significant interaction between expression and identity [$F_{(2, 48)} = 4.15, P = 0.02$]. The interaction was due to differences between the same-identity conditions, but not between the different-identity conditions (Fig. S4). For the same-identity conditions, there was a significant difference between the same- and within-expression conditions [$t_{(19)} = 2.14, P = 0.05$] and the same- and between-expression conditions [$t_{(19)} = 2.51, P = 0.02$]. However, there was no significant difference between the within- and between-expression conditions [$t_{(19)} = 0.69$]. In contrast, for the different-identity conditions, there were no significant difference between the same-expression and either the within-expression [$t_{(19)} = 1.61, P = 1.22$] or between-expression [$t_{(19)} = 0.26$] conditions. There was also no difference in response among the between- and within-expression conditions [$t_{(19)} = 1.44, P = 0.17$].

Discussion

The aim of this study was to determine how facial expressions of emotion are represented in face-selective regions of the human brain. In experiment 1, we found that the pSTS and the amygdala were sensitive to faces that changed in expression and that this sensitivity was largely independent of changes in facial identity. In experiment 2, we morphed between expressions to determine whether the response to expression in the pSTS and the amygdala revealed a categorical or continuous representation. These results clearly show that the pSTS processes facial expressions of emotion using a continuous neural code, whereas the amygdala has a more categorical representation of facial expression.

Our findings offer a compelling alternative to the longstanding controversy about whether facial expressions of emotion are processed using a continuous or categorical code. Behavioral findings consistent with a categorical representation of facial expression are evident when participants report discrete rather than continuous changes in the emotion of faces that are morphed between two expressions (8, 9). Stronger evidence for a categorical representation is seen in the enhanced discrimination of face images that cross a category boundary compared with images that are closer to the prototype expressions (8, 9). Nonetheless, a purely categorical perspective is unable to account for the systematic pattern of confusions that can occur when judging facial expressions (2), and it also has difficulty explaining why morphed expressions that are close to the category prototype are easier to recognize than expressions belonging to the same category but more distant from the prototype (11). So there is evidence to

support both continuous and categorical accounts of facial expression perception (6).

In an attempt to resolve these seemingly conflicting positions, computational models have been developed to explore representations that could underpin both a continuous and a categorical coding of facial expression (22, 23). Our results provide a converging perspective by showing that different regions in the face-processing network can have either a primarily categorical or a primarily continuous representation of facial expression. Of course, the more categorical response in the amygdala compared with the pSTS does not imply that the amygdala is insensitive to changes in facial expression that do not result in a change to the perceived emotion. Indeed, a perceiver needs to be aware of both the category to which a facial expression belongs (its social meaning) and its intensity, and a number of studies have shown that responses in the amygdala can be modulated by changes in the intensity of an emotion (24, 25). Nevertheless, the key finding here is that there is a dissociation between the way facial expressions of emotion are represented in the pSTS and amygdala.

The importance of understanding how facial expressions of emotion are represented in the brain reflects the significance of attributing meaning to stimuli in the environment. When processing signals that are important for survival, perception needs to be prompt and efficient. Categorical representations of expression are optimal for making appropriate physiological responses to threat. The more categorical representation of facial expressions of emotion in the amygdala is consistent with its role in the detection and processing of stimuli pertinent to survival (26, 27). Indeed, neuropsychological studies of patients with amygdala damage have demonstrated impairments in emotion recognition (28–30), which are often accompanied by an attenuated reaction to potential threats (31, 32). Although deficits in emotion recognition following amygdala damage have mostly been reported for the perception of fear, more recent functional neuroimaging studies have provided support for the role of the amygdala in the processing of other emotions (33–35). In the present study, the morphed stimuli were not restricted to those involving fear, so the categorical response shown in this study provides further support for the involvement of the amygdala in processing a range of facially expressed emotions.

However, not all naturally occurring circumstances demand a categorical response, and there are many everyday examples where a continuous representation might be of more value. For example, although there appear to be a small number of basic emotions that seem to be recognized consistently across participants, there are many facial expressions that do not correspond to one or other of these categories. Even with basic emotions, the expressions can actually be quite variable in ways that can signal subtle but important differences (10). Furthermore, judgments of facial expression can be influenced by the context in which they are seen (36). Together, these findings suggest that a more flexible continuous representation is also used in judgments of facial expression. The results from this study suggest that the pSTS could provide a neural substrate for this continuous representation. Our results are consistent with a previous study that used multivariate pattern analysis to show a continuous representation of expression in the pSTS (37). These findings highlight the role of the pSTS in processing moment-to-moment signals important in social communication (38).

An interesting question for further studies concerns how the differences between responses of the pSTS and amygdala are represented at the single-cell level. The blood-oxygen-level–dependent signal measured in fMRI clearly reflects a population response based on the aggregated activity of large numbers of neurons, and there are many ways in which differences in responses across brain regions might therefore be represented in terms of coding by single cells. We offer here a suggestion as to a plausible way in which the population responses shown in our fMRI data might reflect coding by cells in pSTS and amygdala.

Facial expressions are signaled by a complex pattern of underlying muscle movements that create variable degrees of

change in the shapes of facial features, such as the eyes or mouth, the opening or closing of the mouth to show teeth, the positioning of upper and lower teeth, and so on. An obvious hypothesis, then, is that pSTS cells are involved in coding this wide variety of individual feature changes. This hypothesis would be consistent both with the data presented here showing an overall sensitivity of pSTS to any change in expressive facial features and with other fMRI findings that demonstrate pSTS responsiveness both to mouth movements and to eye movements (39). When expressions are perceptually assigned to different emotion categories, however, the underlying feature changes are used simultaneously, so that a particular emotion is recognized through a holistic analysis of a critical combination of expressive features (40). Cells in the amygdala would therefore be expected to have this property of being able to respond to more than isolated features, and it is known, for example, that the amygdala's response to fearful expressions is based on multiple facial cues because it can be driven by different face regions when parts of the face are masked (41).

The same distinction can clearly be seen in computational models of facial expression perception, such as EMPATH (22). EMPATH forms a particularly good example because it is a well-developed, "neural network" model that is able to simulate effects from behavioral studies of facial expression recognition that show continuous or categorical responses in classification accuracy and reaction time. To achieve this goal, EMPATH has layers of processing units that correspond to early stages of visual analysis (Gabor filters, considered as analogous to V1), to principal components (PCs) of variability between facial expressions (as identified by PC analysis of facial expression images), followed by a final classification stage based on the outputs of the PC-responsive layer. There is a clear parallel between coding expression PCs and the fMR properties we report for pSTS and between classifying PC output combinations and the fMR response from the amygdala.

Because of the considerable importance attached to different types of facial information, the most efficient way to analyze this information is thought to involve different neural subcomponents that are optimally tuned for particular types of facial signals (6, 15, 42). Models of face perception suggest that the analysis of invariant cues, such as identity, occurs largely independently of the processing of changeable cues in the face, such as expression. For example, Rotshtein and colleagues used a similar image morphing paradigm to differentiate between regions in the inferior temporal lobe that represent facial identity relative to those regions that are just sensitive to physical properties of the face (43). In this study, we showed that the sensitivity in the response of the pSTS and amygdala to facial expression was largely independent of changes in facial identity. Although the pSTS has been shown to be influenced by identity (14, 16, 44, 45), these results suggest that the neural responses in pSTS and amygdala are primarily driven by changeable aspects of the face, such as expression. In contrast, we found that both the OFA and FFA were sensitive to changes in expression and identity. These findings might be seen as consistent with previous studies that suggest the FFA is involved in judgments of identity and expression (14, 45, 46). However, it is also possible that such results show that the FFA is just sensitive to any structural change in the image.

In conclusion, we found that face-selective regions in the pSTS and amygdala were sensitive to changes in facial expression, independent of changes in facial identity. Using morphed images, we showed that the pSTS has a continuous representation of facial expression, whereas the amygdala has a more categorical representation of facial expression. The continuous representation used by pSTS is consistent with its hypothesized role in processing changeable aspects of faces that are important in social interactions. In contrast, the responses of the amygdala are consistent with its role in the efficient processing of signals that are important to survival.

Methods

Participants and Overall Procedure. All participants were right-handed and had normal or corrected-to-normal vision. All participants provided written consent, and the study was given ethical approval by the York Neuroimaging Centre Ethics Committee.

Each participant took part in one of two experimental scans recording neural responses to conditions of interest and a separate functional localizer scan to provide independent identification of face-selective regions. Visual stimuli (8° × 8°) were back-projected onto a screen located inside the magnetic bore, 57 cm from participants' eyes. Images in all experiments were presented in grayscale by using Neurobehavioral Systems Presentation 14.0. Twenty participants took part in experiment 1 (15 females; mean age, 23), and 25 participants took part in experiment 2 (19 females; mean age, 25).

Experiment 1. Stimuli were presented in blocks, with five images per block. Each face was presented for 1,100 ms and separated by a gray screen presented for 150 ms. Stimulus blocks were separated by a 9-s-fixation gray screen. Each condition was presented 10 times in a counterbalanced order, giving a total of 40 blocks. To ensure that they maintained attention throughout the experiment, participants had to push a button when they detected the presence of a red dot, which was superimposed onto 20% of the images.

Face stimuli were Ekman faces selected from the Young et al. Facial Expressions of Emotion Stimuli and Tests (FEEST) set (47). Five individuals posing five expressions (anger, disgust, fear, happiness, and sadness) were selected, based on the following three main criteria: (*i*) a high recognition rate for all expressions [mean recognition rate in a six-alternative forced-choice experiment: 93% (47)], (*ii*) consistency of the action units (muscle groups) across different individuals posing a particular expression, and (*iii*) visual similarity of the posed expression across individuals. Using these criteria to select the individuals from the FEEST set helped to minimize variations in how the expressions were posed.

Experiment 2. Faces were shown for 700 ms and separated by a 200-ms gray screen. Successive blocks were separated by a 9-s fixation cross. To ensure that participants maintained attention, they had to press a button on detection of a red dot, which was superimposed onto 20% of the images. Images were presented in blocks, with six images per block.

To generate the morphed expression continua for this experiment, we used the five models and five expressions from experiment 1. To control for individual differences in the posed expressions, we used PsychoMorph (48) to generate the average shape of expressions for each emotion category. For each average shape, we then applied the texture from each individual person's face to produce five distinct identities with the same (i.e., equivalently shaped) facial expression. In this way, we ensured that differences between the images used to create experimental stimuli were as tightly dependent on underlying changes in identity or expression as possible.

Expression continua for the experiment were then generated by morphing between different expression images (PsychoMorph). Validation of the morphing procedure was demonstrated in two behavioral experiments. First, we performed an expression-classification experiment, in which 26 participants (19 female; mean age 22) were asked to make a 5AFC for the following expression continua: happy–fear, disgust–happy, sad–disgust, and fear–disgust. Face stimuli were presented for 1,000 ms followed by a 2-s gray screen when participants could make their response. The second behavioral experiment involved a more stringent test of categorical perception using a two-alternative forced-choice discrimination task. Fourteen participants (11 females; mean age 24) were presented with two sequentially presented faces with the same identity and had to judge whether the images were identical or different. Face images were each presented for 900 ms with an inter-stimulus interval of 200 ms.

Localizer Scan. To identify face-selective regions, a localizer scan was conducted after each fMRI experiment. There were five conditions: same-identity faces, different-identity faces, objects, places, and scrambled faces. Images from each condition were presented in a blocked design with five images in each block. Each image was presented for 1 s followed by a 200-ms fixation cross. Blocks were separated by a 9-s gray screen. Each condition was repeated five times in a counterbalanced design. The participants' task was to detect the presence of a red dot that was superimposed on 18% of the images. Face images were taken from the Radboud Face Database (49) and varied in expression. Images of objects and scenes came from a variety of Internet sources.

Imaging Parameters. All imaging experiments were performed by using a GE 3-tesla HD Excite MRI scanner at York Neuroimaging Centre at the University of York. A Magnex head-dedicated gradient insert coil was used in conjunction with a birdcage, radio-frequency coil tuned to 127.4 MHz. A gradient-echo echoplanar imaging (EPI) sequence was used to collect data from 38 contiguous axial slices [time of repetition (TR) = 3, time of echo = 25 ms, field of view = 28 × 28 cm, matrix size = 128 × 128, slice thickness = 4 mm]. These were coregistered onto a T1-weighted anatomical image (1 × 1 × 1 mm) from each participant. To improve registrations, an additional T1-weighted image was taken in the same plane as the EPI slices.

fMRI Analysis. Statistical analysis of the fMRI data was performed by using FEAT (www.fmrib.ox.ac.uk/fsl). The initial 9 s of data from each scan was removed to minimize the effects of magnetic saturation. Motion correction was followed by spatial smoothing (Gaussian, FWHM 6 mm) and temporal high-pass filtering (cutoff, 0.01 Hz). Face-selective regions were individually defined in each individual by using the localizer scan by the average of the following two contrasts: (*i*) same-identity faces > nonface stimuli; and (*ii*) different-identity faces > nonface stimuli. Statistical images were thresholded at $P < 0.001$ (uncorrected). In this way, contiguous clusters of voxels located in the inferior fusiform gyrus, in the posterior occipital cortex, and in the superior temporal lobe of individual participants could be identified as pSTS, OFA, and FFA (Tables S1 and S2 and Fig. S1). A different approach had to be taken to define the amygdala, which is not reliably identified through a functional localizer at the individual level. A face-responsive ROI in the amygdala was therefore defined from the face-selective statistical

map at the group level, thresholded at $P < 0.001$ (uncorrected). This ROI in the amygdala was then transformed into the individual MRI space for each participant. The time course of response in the amygdala ROI was then evaluated for each participant to ensure that it responded more to faces than to nonface stimuli. In addition to these functional criteria, we were able to define the amygdala based on anatomy. Despite the difference in the way that the amygdala was defined, Fig. S5 shows that the face-selective voxels that are located in the amygdala show a corresponding face selectivity to the other ROI. In all other respects, the data were processed in exactly the same way for all ROIs.

For each experimental scan, the time series of the filtered MR data from each voxel within a ROI was converted from units of image intensity to percentage signal change. All voxels in a given ROI were then averaged to give a single time series for each ROI in each participant. Individual stimulus blocks were normalized by subtracting every time point by the zero point for that stimulus block. The normalized data were then averaged to obtain the mean time course for each stimulus condition. There was no main effect of hemisphere in experiment 1 [$F_{(1, 16)} = 0.39$] and experiment 2 [$F_{(1, 17)} = 2.96$, $P = 0.13$], so data from the left and right hemispheres were combined for all ROIs. Repeated-measures ANOVA was used to determine significant differences in the peak response to each stimulus condition in each experiment. The peak response was taken as an average of the TR 2 (6 s) and TR 3 (9 s) following stimulus onset.

1. Ekman P, Friesen VW (1978) *Facial Action Coding System: A Technique for the Measurement of Facial Movement* (Consulting Psychologists Press, Palo Alto, CA).
2. Woodworth RS, Schlosberg H (1954) *Experimental Psychology* (Henry Holt, New York), Revised Ed.
3. Russell JA (1980) A circumplex model of affect. *J Pers Soc Psychol* 39:1161–1178.
4. Darwin C (1998) *The Expression of the Emotions in Man and Animals*, ed Ekman P (HarperCollins, London).
5. Ekman P (1999) Facial expressions. *The Handbook of Cognition and Emotion*, eds Dalgleish T, Power T (John Wiley & Sons, Sussex, UK), pp 301–320.
6. Bruce V, Young AW (2012) *Face Perception* (Psychology Press, Hove, England).
7. Ekman P (1972) Universals and cultural differences in facial expressions of emotion. *Nebraska Symposium on Motivation, 1971*, ed Cole J (Univ of Nebraska Press, Lincoln), pp 207–283.
8. Etcoff NL, Magee JJ (1992) Categorical perception of facial expressions. *Cognition* 44(3):227–240.
9. Calder AJ, Young AW, Perrett DI, Etcoff NL, Rowland D (1996) Categorical perception of morphed facial expressions. *Vis Cogn* 3:81–117.
10. Rozin P, Lowery L, Ebert R (1994) Varieties of disgust faces and the structure of disgust. *J Pers Soc Psychol* 66(5):870–881.
11. Calder AJ, Young AW, Rowland D, Perrett DI (1997) Computer-enhanced emotion in facial expressions. *Proc Biol Sci* 264(1383):919–925.
12. Young AW, et al. (1997) Facial expression megamix: Tests of dimensional and category accounts of emotion recognition. *Cognition* 63(3):271–313.
13. Breiter HC, et al. (1996) Response and habituation of the human amygdala during visual processing of facial expression. *Neuron* 17(5):875–887.
14. Fox CJ, Moon SY, Iaria G, Barton JJS (2009) The correlates of subjective perception of identity and expression in the face network: An fMRI adaptation study. *Neuroimage* 44(2):569–580.
15. Haxby JV, Hoffman EA, Gobbini MI (2000) The distributed human neural system for face perception. *Trends Cogn Sci* 4(6):223–233.
16. Winston JS, Henson RNA, Fine-Goulden MR, Dolan RJ (2004) fMRI-adaptation reveals dissociable neural representations of identity and expression in face perception. *J Neurophysiol* 92(3):1830–1839.
17. Puce A, Allison T, Bentin S, Gore JC, McCarthy G (1998) Temporal cortex activation in humans viewing eye and mouth movements. *J Neurosci* 18(6):2188–2199.
18. Pelphrey KA, Viola RJ, McCarthy G (2004) When strangers pass: Processing of mutual and averted social gaze in the superior temporal sulcus. *Psychol Sci* 15(9):598–603.
19. Calder AJ, et al. (2007) Separate coding of different gaze directions in the superior temporal sulcus and inferior parietal lobule. *Curr Biol* 17(1):20–25.
20. Engell AD, Haxby JV (2007) Facial expression and gaze-direction in human superior temporal sulcus. *Neuropsychologia* 45(14):3234–3241.
21. Baseler H, Harris RJ, Young AW, Andrews TJ (2012) Neural responses to expression and gaze in the posterior superior temporal sulcus interact with facial identity. *Cereb Cortex*, 10.1093/cercor/bhs360.
22. Dailey MN, Cottrell GW, Padgett C, Adolphs R (2002) EMPATH: A neural network that categorizes facial expressions. *J Cogn Neurosci* 14(8):1158–1173.
23. Martinez AM, Du S (2012) A model of the perception of facial expressions of emotion by humans: Research overview and perspectives. *J Mach Learn Res* 13:1589–1608.
24. Morris JS, et al. (1996) A differential neural response in the human amygdala to fearful and happy facial expressions. *Nature* 383(6603):812–815.
25. Thielscher A, Pessoa L (2007) Neural correlates of perceptual choice and decision making during fear-disgust discrimination. *J Neurosci* 27(11):2908–2917.
26. Adolphs R, et al. (1999) Recognition of facial emotion in nine individuals with bilateral amygdala damage. *Neuropsychologia* 37(10):1111–1117.
27. Sander D, Grafman J, Zalla T (2003) The human amygdala: An evolved system for relevance detection. *Rev Neurosci* 14(4):303–316.
28. Adolphs R, Tranel D, Damasio H, Damasio A (1994) Impaired recognition of emotion in facial expressions following bilateral damage to the human amygdala. *Nature* 372(6507):669–672.
29. Anderson AK, Phelps EA (2000) Expression without recognition: Contributions of the human amygdala to emotional communication. *Psychol Sci* 11(2):106–111.
30. Young AW, et al. (1995) Face processing impairments after amygdalotomy. *Brain* 118(Pt 1):15–24.
31. Sprengelmeyer R, et al. (1999) Knowing no fear. *Proc Biol Sci* 266(1437):2451–2456.
32. Feinstein JS, Adolphs R, Damasio AR, Tranel D (2011) The human amygdala and the induction and experience of fear. *Curr Biol* 21(1):34–38.
33. Phan KL, Wager T, Taylor SF, Liberzon I (2002) Functional neuroanatomy of emotion: A meta-analysis of emotion activation studies in PET and fMRI. *Neuroimage* 16(2):331–348.
34. Winston JS, O'Doherty J, Dolan RJ (2003) Common and distinct neural responses during direct and incidental processing of multiple facial emotions. *Neuroimage* 20(1):84–97.
35. Sergerie K, Chochol C, Armony JL (2008) The role of the amygdala in emotional processing: A quantitative meta-analysis of functional neuroimaging studies. *Neurosci Biobehav Rev* 32(4):811–830.
36. Russel JA, Fehr B (1987) Relativity in the perception of emotion in facial expressions. *J Exp Psychol* 116:233–237.
37. Said CP, Moore CD, Norman KA, Haxby JV, Todorov A (2010) Graded representations of emotional expressions in the left superior temporal sulcus. *Front Syst Neurosci* 4:6.
38. Allison T, Puce A, McCarthy G (2000) Social perception from visual cues: Role of the STS region. *Trends Cogn Sci* 4(7):267–278.
39. Pelphrey KA, Morris JP, Michelich CR, Allison T, McCarthy G (2005) Functional anatomy of biological motion perception in posterior temporal cortex: An FMRI study of eye, mouth and hand movements. *Cereb Cortex* 15(12):1866–1876.
40. Calder AJ, Young AW, Keane J, Dean M (2000) Configural information in facial expression perception. *J Exp Psychol Hum Percept Perform* 26(2):527–551.
41. Asghar AUR (2008) An amygdala response to fearful faces with covered eyes. *Neuropsychologia* 46(9):2364–2370.
42. Bruce V, Young A (1986) Understanding face recognition. *Br J Psychol* 77(Pt 3):305–327.
43. Rotshtein P, Henson RNA, Treves A, Driver J, Dolan RJ (2005) Morphing Marilyn into Maggie dissociates physical and identity face representations in the brain. *Nat Neurosci* 8(1):107–113.
44. Andrews TJ, Ewbank MP (2004) Distinct representations for facial identity and changeable aspects of faces in the human temporal lobe. *Neuroimage* 23(3):905–913.
45. Cohen Kadosh K, Henson RNA, Cohen Kadosh R, Johnson MH, Dick F (2010) Task-dependent activation of face-sensitive cortex: An fMRI adaptation study. *J Cogn Neurosci* 22(5):903–917.
46. Ganel T, Valyear KF, Goshen-Gottstein Y, Goodale MA (2005) The involvement of the "fusiform face area" in processing facial expression. *Neuropsychologia* 43(11):1645–1654.
47. Young AW, Perrett DI, Calder AJ, Sprengelmeyer R, Ekman P (2002) *Facial Expressions of Emotion Stimuli and Tests (FEEST)* (Thames Valley Test Company, Bury St Edmunds, England).
48. Tiddeman B, Burt DM, Perrett DI (2001) Computer graphics in facial perception research. *IEEE Comput Graph Appl* 21:42–50.
49. Langner O, et al. (2010) Presentation and validation of the Radboud Faces Database. *Cogn Emotion* 24:1377–1388.

PSYCHOLOGICAL AND COGNITIVE SCIENCES