

# Contrasting Specializations for Facial Motion within the Macaque Face-Processing System

Clark Fisher<sup>1,\*</sup> and Winrich A. Freiwald<sup>1,\*</sup>

<sup>1</sup>Laboratory of Neural Systems, The Rockefeller University, 1230 York Avenue, New York, NY 10065, USA

## Summary

Facial motion transmits rich and ethologically vital information [1, 2], but how the brain interprets this complex signal is poorly understood. Facial form is analyzed by anatomically distinct face patches in the macaque brain [3, 4], and facial motion activates these patches and surrounding areas [5, 6]. Yet, it is not known whether facial motion is processed by its own distinct and specialized neural machinery, and if so, what that machinery's organization might be. To address these questions, we used fMRI to monitor the brain activity of macaque monkeys while they viewed low- and high-level motion and form stimuli. We found that, beyond classical motion areas and the known face patch system, moving faces recruited a heretofore unrecognized face patch. Although all face patches displayed distinctive selectivity for face motion over object motion, only two face patches preferred naturally moving faces, while three others preferred randomized, rapidly varying sequences of facial form. This functional divide was anatomically specific, segregating dorsal from ventral face patches, thereby revealing a new organizational principle of the macaque face-processing system.

## Results

### Face Motion Activates a Diverse Set of Functionally Specific Areas

Face motion activates a large expanse of cortex in and around the superior temporal sulcus (STS) [5, 6]. The degree to which this merely reflects underlying sensitivity to general motion or face form remains unclear. We examined the functional basis of this activation by mapping it alongside regions specialized for general low-level motion and face form, using high-resolution, contrast-enhanced fMRI to monitor brain activity in four alert rhesus macaque monkeys (M1–M4) during visual stimulus presentation (Figures S1A and S2 and the [Supplemental Experimental Procedures](#) available online). We used this same basic technique throughout this study. The resulting functional maps (Figures S2A and S2B) revealed motion areas including MT, MSTv, FST, and LST [7, 8], and face patches PL (posterior lateral), ML (middle lateral), MF (middle fundus), AL (anterior lateral), and AF (anterior fundus) [3]. Motion areas and face patches in the STS fundus, despite their proximity, remained spatially disjunct (Figures S2C and S2D). Face motion activated some face patches, all identified motion areas, and further outlying areas (Figure 1, left column). Nonface object motion (Figure 1, center column) did not activate any face patches, but activated all identified motion areas and a subset

of the aforementioned outlying areas (Figure 1, right column). Outlying areas responsive to both face- and nonface motion likely represent specializations for forms in motion [9] that lack a specificity for faces. Importantly, we also found outlying areas that were recruited by face motion, but neither object nor general motion (Figure 1, white asterisks). These maps show that responses to face motion extend throughout the motion-sensitive STS, into at least a subset of face patches, and, intriguingly, beyond the classical face patch system and motion areas.

### A Novel Face Patch Responds to Moving Faces

To extend beyond the classical face patch system and map areas that may be attuned to the motion of faces [10, 11], we contrasted fMRI responses to movies of faces with responses to movies of articulated toys (Figure S1C). This dynamic localizer (Figure 2A) activated all of the earlier-identified face patches (Figure S2B) and additional parts of the STS's dorsal bank, including many of the areas that had been selectively recruited by face motion (Figure 1, white asterisks). These new dorsal activations included scattered points of face selectivity that varied from individual to individual and, importantly, one area of selectivity at a consistent location in every subject and hemisphere. This area was located anterodorsal to face patches ML and MF (Figure 2A), spatially distinct from both (Figures 2B, 2C, and S3). We call this new area the middle dorsal face patch (MD). Thus, the pairing of face form and motion reliably recruits six face-specific patches around the STS: PL, ML, and AL along its ventral lip, and MF, AF, and the just-recognized MD in its fundus and dorsal bank.

### All STS Face Patches Possess a Distinctive Selectivity for Face Motion

The preference for moving faces over moving objects in these six face patches could result from two different specializations: selectivity either for face form or for face motion. In fact, all face patches demonstrated similar degrees of selectivity for facial form (Figure 2D) and a preference for facial motion (Figure 2E). The facial motion preference was more pronounced in the patches along the fundus and dorsal bank of the STS. Responses to nonface object motion (Figure 2F) were smaller than responses to facial motion throughout. Consistent with this, the interaction between shape category (face versus object) and motion (moving versus static) was significant in all STS face patches (Figure 2G). Thus, all face patches exhibit a response to motion that is face specific. Two neighboring control areas, an object-selective STS region that responded more to moving toys than moving faces (referred to as the “toy patch”) and motion area LST [7], were sensitive to both face and object motion to a similar extent (Figures 2E–2G). The observed form-specific motion selectivity of the face patches is therefore not due to an imbalance of low-level motion energy across stimuli, a conclusion further supported by balanced activation of general motion areas (Figure S4). Thus, selectivity for both the form and motion of faces characterizes all STS face patches, but not the STS at large.

\*Correspondence: [cfisher@rockefeller.edu](mailto:cfisher@rockefeller.edu) (C.F.), [wfreiwald@rockefeller.edu](mailto:wfreiwald@rockefeller.edu) (W.A.F.)

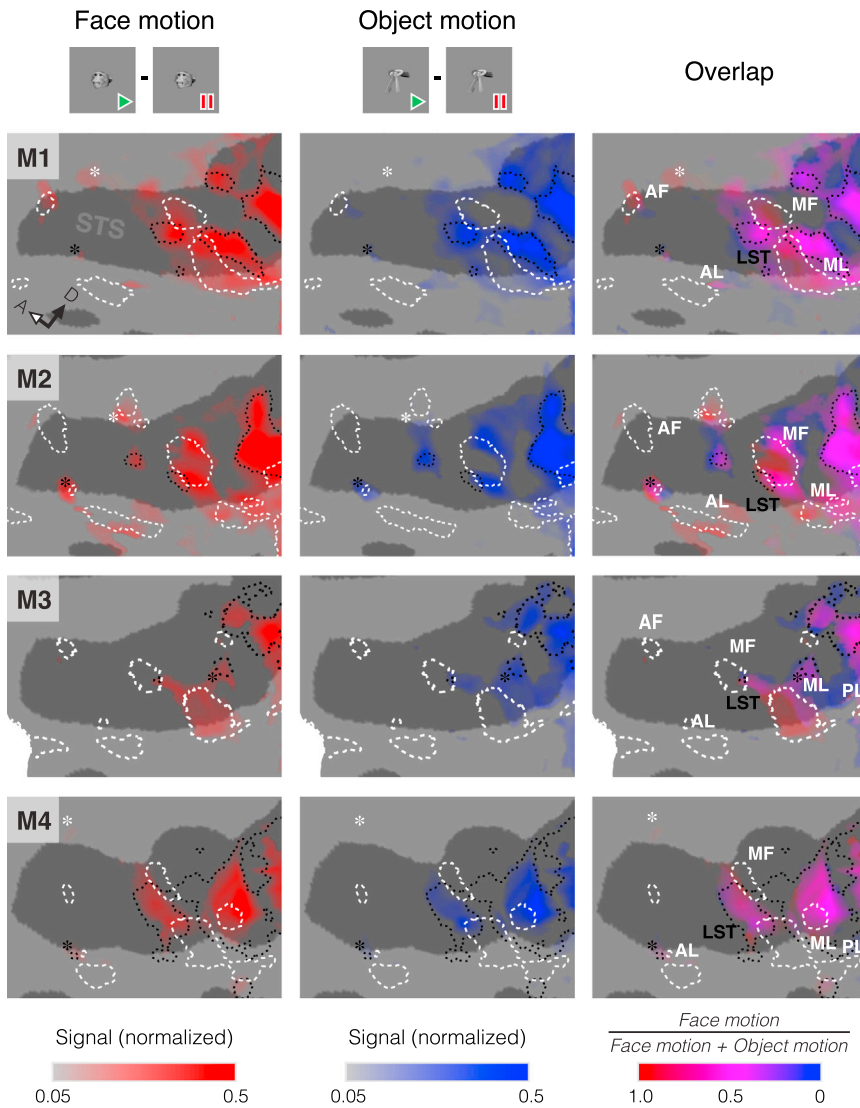


Figure 1. Selectivities for Motion Carried by Faces or Nonface Objects along the Macaque STS

Regions responding to face motion (left column, red; natural face movies – face pictures) or non-face object motion (middle column, blue; natural object movies – object pictures), and the relative strength of these contrasts (right column), in the left hemisphere of each subject. Opacity reflects the contrast strength (normalized signal change). These data are presented on a flattened cortical model of the area surrounding the STS, with dark gray regions representing sulci and light gray regions representing gyri (as in Figure S2B). Dashed white lines outline areas of static face selectivity and dotted black lines outline areas of low-level motion selectivity, both measured in independent experiments (Figure S2). Similarly, white labels indicate face patches and black labels indicate motion areas. Black asterisks highlight areas responding to face and object motion outside of recognized motion-processing areas. White asterisks highlight areas more activated by face motion than object motion outside of known face patches. For orientation, the white-filled arrow points anteriorly and the black arrow points dorsally. Signal change in maps is normalized per subject and thresholded at a false discovery rate of  $q < 0.01$ . See also Figure S2.

preferentially to facial shapes undergoing rapid, even random, transitions.

The divergent responses of face patches to natural versus jumbled motion did not extend to nonface objects: no patch preferred jumbled object movies to natural ones (Figure 3D). Furthermore, the two control regions responded more to natural object motion (compared to jumbled object motion) than to natural facial motion (compared to jumbled facial motion). As a result,

face patches PL, ML, and MF and control area LST showed significant interactions between motion quality (natural or jumbled) and form (face or object; Figure 3E). Thus, while natural motion improved localization of an extended face-processing system (Figure 2A), and all constituent areas of this system were selective for an interaction of face form and motion (Figure 2G), this shared selectivity arose from two different specializations: the dorsal face patches (MD and AF) genuinely represent natural facial motion, while the ventral face patches (PL, ML, and AL) appear to prefer rapidly changing facial pose, regardless of kinematic meaning.

## Discussion

From just a glance at a face, we gather an abundance of social information [12]. Set in motion, the face comes alive, not only augmenting this knowledge [13, 14] but also posing a challenge for the neural systems that must now interpret an evolving subject [15]. The current study aimed to identify the neural machinery that navigates these intertwined opportunities and challenges of facial motion, leveraging a model system that is similar to the human face-processing system

## Natural Face Motion Selectivity Divides the STS Face Patch System

We now know that all STS face patches are selective for facial motion (Figure 2G). But does activity within these areas represent *natural* facial motion, or is it simply a response to all updates in face pose, natural or unnatural? We addressed this question by challenging the face-processing system with two stimulus sets that were identical in static content and frame rate, but differed in motion *quality*: the normal, “natural” movies used earlier in the study (Figure 1), and “jumbled” versions of the same movies, where frames were presented in a random order (Figures 3A and S1C). We found that dorsal face patches MD and AF showed a significantly greater response to natural movies of faces (Figure 3B). In contrast, ventral face patches PL, ML, and AL not only failed to respond more to natural movies, but also, surprisingly, their responses were significantly stronger for jumbled movies. MF, positioned between MD and ML, showed no significant preference for either movie type. Thus, the face patch system is fundamentally differentiated along a ventrodorsal axis (Figure 3C): the dorsal portion responds preferentially to natural face movements and the ventral portion responds

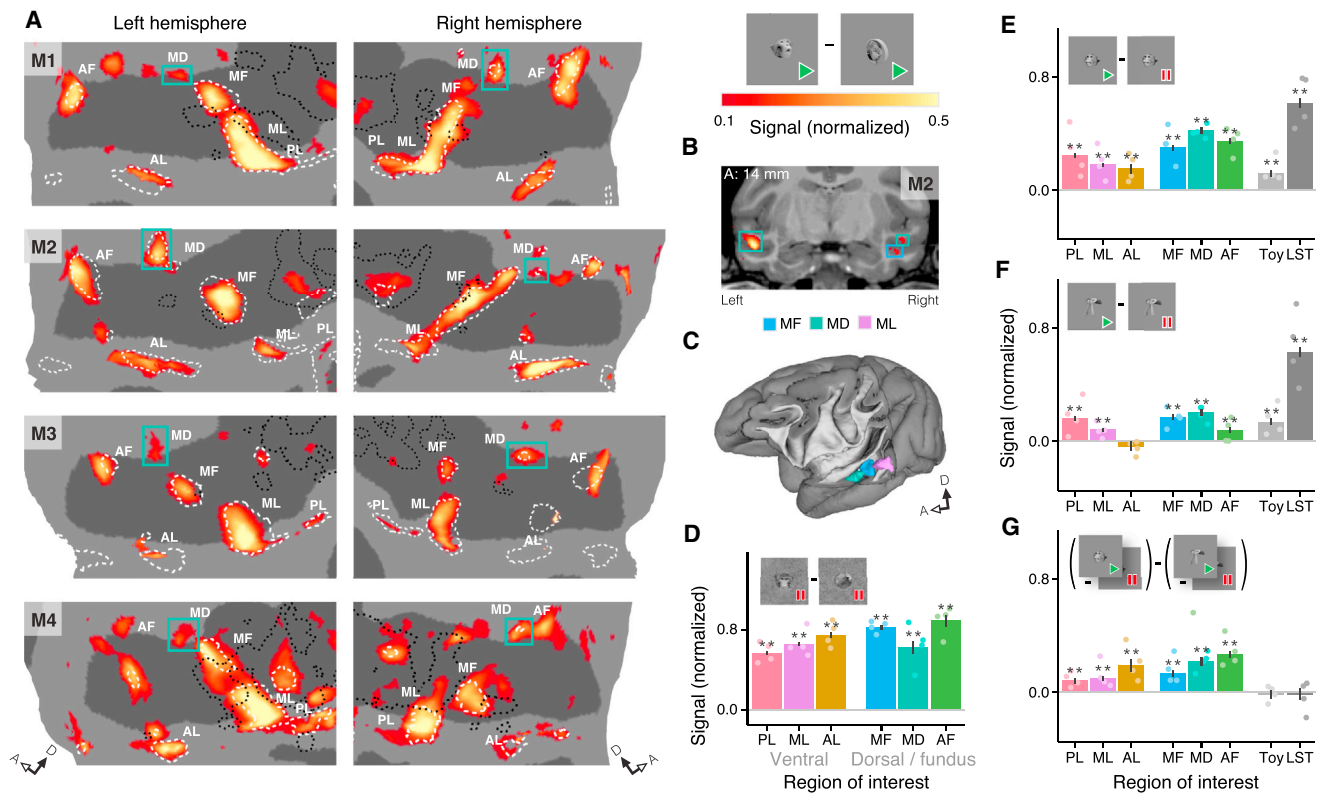


Figure 2. Responses to Complex Motion within an Extended Face Patch System

(A) Left: dynamic face selectivity (localizer face movies – localizer object movies; Figure S1C) on flattened maps of the STS of each hemisphere in four subjects. Green boxes highlight the newly described MD face patch, so-called because it is in the dorsal bank of the STS and neighbors other middle face patches ML and MF. Dashed white lines outline static face selectivity and dotted black lines outline low-level motion selectivity (as in Figure 1). Right: signal strength color map and schematic of contrast.

(B) Coronal slice from M2, showing position of MD and its separation from MF. The right side of the brain is on the right side of the page. The anterior-posterior stereotaxic coordinate is taken relative to the interaural line.

(C) Volumetric model of M2's left hemisphere showing the relative locations of ML (purple), MF (blue), and MD (green).

(D) Plot of static face selectivity (faces – fruits and vegetables) within the six face patch regions of interest (ROIs) that were defined with the dynamic face selectivity localizer (A).

(E and F) Responses to face motion (natural face movies – face pictures) (E) and object motion (natural object movies – object pictures) (F) in the face patches, toy patch, and LST.

(G) Strength of interaction between responses to form and motion: (natural face movies – face pictures) – (natural object movies – object pictures).

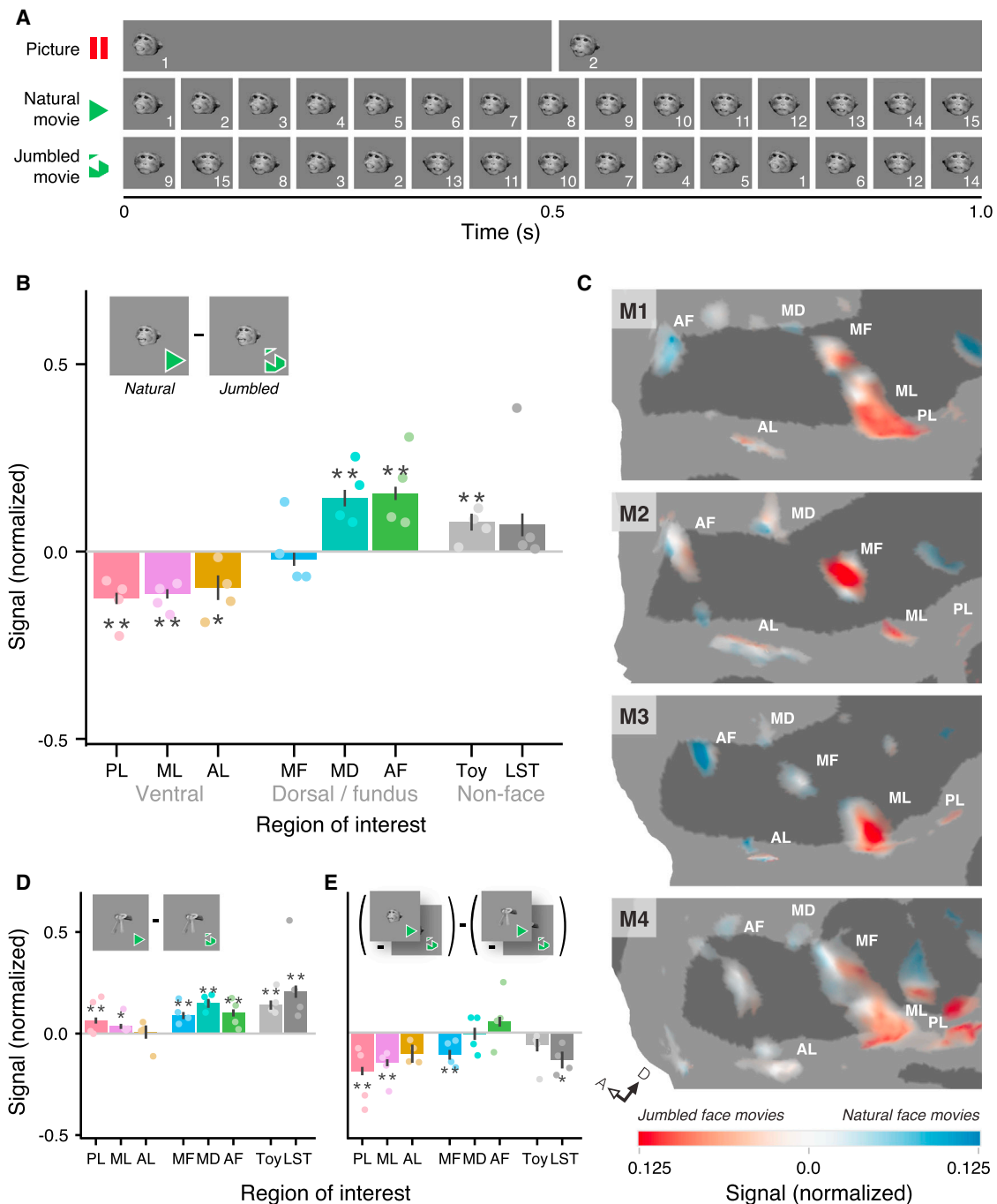
\* $p < 0.05$  and \*\* $p < 0.01$ , corrected using Holm-Bonferroni method for 30 tests (six ROIs  $\times$  one measure + eight ROIs  $\times$  three measures). Dots on bar plots represent the values for individual subjects. Error bars represent SE. Signal change in bar plots is normalized per ROI. Signal change in maps is normalized per subject and thresholded at a false discovery rate of  $q < 0.01$ . The raw data analyzed in (E)–(G) are the same data plotted in Figure 1.

See also Figures S3 and S4.

[3, 16]; remains highly reproducible across subjects [3]; and enables mechanistic exploration of the computations underlying face recognition [4]. The specialized areas that we recruited with naturally moving faces likely mark a key component of the machinery for dynamic face recognition.

The architecture of face motion processing revealed here includes areas selective for low- and high-level motion [5], face form [6], and natural facial motion. These areas all neighbor each other but remain spatially distinct. This picture of a functionally heterogeneous mosaic represents a fundamental departure from earlier fMRI studies [5, 6] that suggested that any motion responsiveness found in dorsal face patches [6] was a by-product of these areas overlapping a generally motion-responsive region. Our results reflect a different reality: while some STS regions are broadly motion sensitive—responding similarly to face motion and nonface motion—neighboring areas specifically process face motion.

One such area is MD, a newly described face patch in the upper bank of the STS (Figures 2A–2C and S3). While MD is occasionally evident when static stimuli are used for mapping (similar to aMF, the recently reported “anterior MF” face patch [17]), dynamic stimuli allowed us to locate this area in all eight hemispheres that we studied. This is reminiscent of the human face area in the posterior superior temporal sulcus (pSTS face area; sometimes called STS face area [16]), a region critical for processing moving faces [18] that is likewise identified sporadically with static stimuli but reliably with dynamic ones [6, 10] and shows selectivity for natural face motion [19]. Interestingly, the human pSTS face area does not appear to be strongly connected to the ventrally located fusiform and occipital face areas [20]. Similarly, in macaque monkeys, when connectivity of face patches was mapped, no strong projections to the location of MD were reported [21]. Furthermore, anatomically variable activations by faces are found anterior to both MD (this study, [17]) and human pSTS [11]. One plausible scenario for



**Figure 3. Preferential Responses to Natural or Disordered Face Motion within the Face Patch System**

(A) Schematics of stimuli used for analyses of natural motion selectivity. For faces and nonface objects, picture, natural movie, and jumbled movie stimuli were derived from the same 60 frames per s (fps) source videos. Each video was downsampled to 2 fps in the picture condition and 15 fps in the movie conditions. By randomizing the order of each natural movie's frames, a matched jumbled movie was created. Each exemplar stimulus lasted 3 s; a 1 s period is shown here for demonstration.

(B) Preference for natural face motion over jumbled face motion across six face patches and two control regions. Dorsal patches MD and AF show a significant preference for natural face motion, while, conversely, ventral patches PL, ML, and AL significantly prefer the rapidly changing jumbled face movies. (C) Preference for either natural face motion (red) or jumbled face motion (blue), as calculated in (B), across face-selective cortex. Opacity reflects strength of face selectivity (Figure 2A).

(D) Preference for natural object movies over jumbled object movies.

(E) Strength of interaction between form (face or object) and frame ordering (natural or jumbled): (natural face movies – jumbled face movies) – (natural object movies – jumbled object movies).

\* $p < 0.05$  and \*\* $p < 0.01$ , corrected using Holm-Bonferroni method for 24 tests (eight ROIs  $\times$  three measures). Dots on bar plots represent the values for individual subjects. Error bars represent SE. Signal change in bar plots is normalized per ROI. Signal change in maps is normalized per subject.

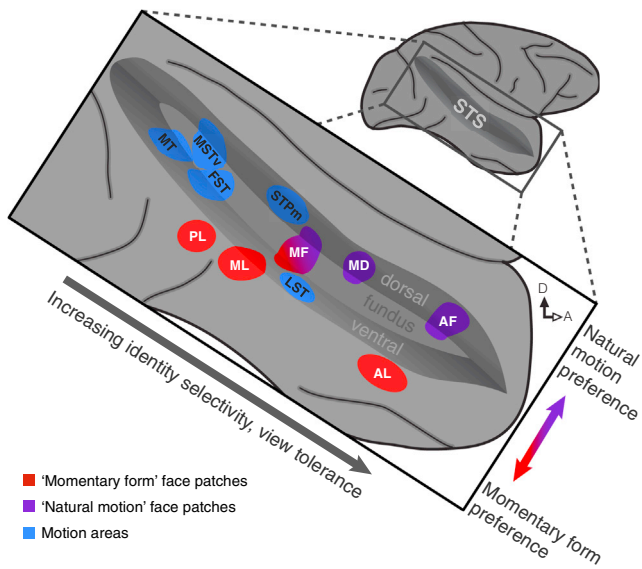


Figure 4. Model of Face Motion and Face Form Processing along the Macaque Temporal Lobe

Functional specificity of face patches is organized along two main anatomical axes. From posterior to anterior, face patches show increasing identity selectivity and increasing tolerance to viewing condition [4]. Along the dorsoventral axis, face patches show differential selectivity for natural motion, with “dynamic” dorsal patches (purple) responding to natural motion and “static” ventral patches (red) responding more to rapidly varying face stimuli. Face motion activates all of these patches as well as motion-processing areas (blue), which are selective for neither momentary face form nor natural face motion.

this variability is that these anterior regions represent a variety of social signals of diverse complexity [22, 23] and that they are only partially and erratically activated by faces. Thus, functional specialization, connectivity, and relative location indicate that MD might be the macaque homolog of the human pSTS face area and could therefore be critical for establishing general homology between face-processing systems of humans and macaques [3, 16].

We found a new functional differentiation within the macaque face-processing system wherein dorsal patches preferred naturally moving faces, while ventral patches (to our surprise) preferred random transitions in face pose (Figures 3B and 3C). This reveals a novel dimension of the cortical representation of faces and marks, to our knowledge, the first time that fMRI has revealed an overt functional dissociation—where different areas have significant and opposing selectivities—within the macaque face-processing system.

This preference for natural facial motion suggests that cells in MD and AF, beyond selectivity for static facial form (Figure 2D; see [24]), also exhibit selectivity for the kinematics of naturally moving faces. Some neurons in these patches may fire only in response to a specific sequence of poses, a mechanism that has been proposed for the neural coding of biological motion [25, 26]. On the other hand, the apparent selectivity of ventral face patches PL, ML, and AL for randomized face motion is unlikely to reflect a genuine selectivity for specific sequences of facial pose. Rather, this preference may reflect purely shape-selective face neurons that adapt quickly [27], respond less to expected stimuli [28, 29], or show a combination of these effects [30, 31]. Thus, a predictive coding scheme, where deviations from expectation drive neural activity [32,

33], could underlie processing in the ventral patches. While predictive coding models generally assume predictions from later processing levels inform earlier processing levels (e.g., [32]), our discovery of qualitatively distinct representations of facial motion within the face patch system allows an alternative hypothesis to be explored: dynamic face representations in dorsal face patches might generate predictions of momentary features that are communicated to ventral patches through lateral connections [21].

While our use of jumbled frames as a control revealed a functional divide within the face patch system, jumbling is a coarse manipulation that introduces discontinuities into continuous motion and interrupts the possible expectation of preserved stimulus identity. This study, therefore, speaks specifically to functional specializations for continuous face motion. A recent experiment demonstrated that certain human face areas respond differentially to movies of facial expressions played either forward in time (a continuous, biologically plausible motion) or backward (a continuous but implausible motion) [34]. A similar comparison in monkeys might deepen our understanding, showing further motion specialization within the face patches or refining the mechanistic understanding of the division we describe.

By integrating our results with the findings of earlier studies, we can develop a picture of how face form and motion processing are arranged in the macaque temporal lobe (Figure 4). Within and around the STS, face patches and general motion areas adjoin each other, but are anatomically distinct. The face patches are differentiated along two axes. As information flows from posterior to anterior, face patches show increased form specificity and view tolerance [4], consistent with general trends in the temporal lobe [35]. Along the ventral-to-dorsal axis, there is a functional transition that reflects a likely selectivity for momentary facial form in the ventral patches and for continuous facial motion in the dorsal ones. This picture is compatible with influential “division of labor” face recognition models (e.g., [36]), particularly those that posit a separation of static features (such as identity) from dynamic ones (such as expression) [37]. In fact, our findings present the best evidence yet of such a division of labor between identifiable nodes in the macaque brain, opening the door to further characterization of putative static and dynamic streams by electrophysiological and causal approaches. This could ultimately elucidate how the myriad signals conveyed by faces are given meaning by the brain [38, 39] at neuron and network levels. In this way, the specializations for facial motion within the areas described here provide a concrete anatomical framework for investigating both the computations that extract and abstract from facial dynamics and, more generally, the interrelated neural representations of form and motion.

#### Supplemental Information

Supplemental Information includes Supplemental Experimental Procedures and four figures and can be found with this article online at <http://dx.doi.org/10.1016/j.cub.2014.11.038>.

#### Author Contributions

C.F. and W.A.F. designed the experiment. C.F. performed the experiment and analyzed the data. C.F. and W.A.F. wrote the manuscript.

#### Acknowledgments

We thank Akinori Ebihara, Rizwan Huq, Sofia Landi, Srivatsun Sadagopan, Caspar Schwiedrzik, Stephen Shepherd, Julia Sliwa, and Wilbert Zarco for

