Title: Delayed suppression normalizes face identity responses in the primate brain

Authors: Kenji W. Koyano[1,*,†], Adam P. Jones[1,†], David B. T. McMahon[1,2], Elena N. Waidmann[1], Brian E. Russ[1,3], David A. Leopold[1,4,*]

Affiliations:

[1] Section on Cognitive Neurophysiology and Imaging, National Institute of Mental Health, 49 Convent Dr., Bethesda, MD 20892, USA.

[2] Neuronal Networks Section, National Eye Institute, 49 Convent Dr., Bethesda, MD 20892, USA.

[3] Center for Biomedical Imaging and Neuromodulation, Nathan S. Kline Institute for Psychiatric Research, 140 Old Orangeburg Road, Orangeburg, New York, 10962, USA.

[4] Neurophysiology Imaging Facility, National Institute of Mental Health, National Institute of Neurological Disorders and Stroke, National Eye Institute, 49 Convent Dr., Bethesda, MD 20892, USA.

[*]Corresponding author. Email: leopoldd@mail.nih.gov or kenji.koyano@nih.gov

[†] These authors contributed equally to this work.

Abstract: The primate brain is specialized for social visual perception. Previous work indicates that recognition draws upon an internal comparison between a viewed face and an internally stored average face. Here we demonstrate that this comparison takes the form of a delayed, dynamic suppression of face averageness among single neurons. In three macaque face patches, spiking responses to low-identity morphed faces met with a synchronous attenuation starting approximately 200 ms after onset. Analysis showed that a late-emerging V-shaped identity tuning was sometimes superimposed on linear ramp tuning and could not be ascribed to repetition suppression. The results indicate that the brain's analysis of faces is enhanced through predictive normalization of identity, which increases sensitivity among face-selective neurons to distinctive facial features known to drive recognition.

One Sentence Summary: Dynamic suppression underlies identity normalization in macaque face patch neurons.

Main Text:

Humans and other primates rely primarily on vision to extract social information, such as identity and emotional state, from their conspecifics (*1*). Primates are highly adapted for face perception, which is reflected in multiple fMRI-defined face patches observed in the inferotemporal cortex of humans, macaques, and marmosets (*2-4*). The recognition of facial identity is a difficult visual problem that requires extracting the distinctive features of individual faces. Human psychophysical experiments indicate that the brain approaches face recognition by employing an internally stored reference of face averageness, which acts to accentuate feature combinations specific to an individual (*5-8*). This framework has been formalized as a "face-space" centered on the average face. In this conceptual space, radial trajectories emanating outward from the center are axes of caricaturization that indicate graded levels of individual identity (*9*).

While much evidence supports the importance of the average face in the perception of face identity, researchers are divided on how this computation may be implemented in the brain (*10-16*). A previous electrophysiological study demonstrated that neurons in the macaque inferotemporal cortex showed graded responses to face identity in a manner that resembled the structure of face space, eliciting minimal response to the average face and then systematically increasing their firing rate to more extreme identities (*10*). Human experiments have shown similar patterns in fMRI responses to faces (*11*), as well as to human voices (*12*) and learned shapes (*13*). More recent single unit recordings from macaque face patches also found graded responses to

extreme faces, but, unlike the previous studies, did not identify a particular role for average facial features (*14, 15*). The basis of this discrepancy is an open and important question.

An interesting recent perspective on face processing places it in the context of predictive coding and links norm-based coding to hierarchical error signals in the visual system (*17*). In this study, single neurons recorded in the inferior temporal cortex altered their late-phase responses to faces according to the typicality of facial feature configurations. Might the brain's analysis of facial identity be construed as a type of predictive coding where a stable representation of face averageness serves as the prediction? Mechanistically, this might entail the brain's active discounting of average facial structural components in order to achieve normalization in the domain of facial identity. Like other examples of normalization in the nervous system, this operation would serve to adjust sensory neurons' dynamic range to be most sensitive to dimensions that are important for behavior (*18*). Here we report that single neurons in three fMRI-defined macaque face patches dynamically express such discounting of face averageness through a delayed and synchronous suppression of neural spiking to low-identity face components, resulting in identity normalization.

We quantitively assessed spiking responses to monkey faces morphed along multiple continuous identity trajectories spanning the center of face space (Fig. 1A). Each trajectory included diminished, full, and exaggerated identity morphs for a given individual (i.e. positive identities), in addition to so-called "anti-faces" (i.e. negative identities). Importantly, anti-faces were generated by extrapolating the morphing transformation beyond the center of face space, such that the features of an anti-face were nominally opposite to those of the original face (*19*). We recorded spiking responses to 145 photorealistic monkey faces (Fig. 1A-B, Fig. S1) from a total of 296 neurons in three face patches (middle lateral, ML; anterior fundus, AF; anterior medial,

AM) of six rhesus macaques (Fig. 1C). We restricted analysis to 186 neurons (103 in AM, 67 in AF and 16 in ML) that showed selectivity for monkey facial identity ($p < 0.05$, two-way ANOVA).

Across the population, most neurons exhibited a conspicuous dip in response magnitude centered on the average face, resulting in a V-shaped identity tuning profile (Fig. 1D, 2A). A smaller subset of cells showed the opposite profile, with a peak instead of a dip (Fig. S2B). For a given cell, different identity trajectories could exhibit either a slope reversal at the average face (V- or inverse-V-shape), an inflection point (knee-shape), or neither (linear ramp tuning) with different neurons showing varied combinations of identity tuning profiles (Fig S2). For the majority of neurons, averaging across all twelve identity trajectories resulted in clear V-shaped tuning (thick black lines in Fig. 2B and Fig. S2), which was evident across the population of all face-selective neurons (Fig. 2C, Fig. S3). This tuning profile could not arise as the sum or average of linear ramp responses to different identities, but instead depended upon the slope discontinuity at low identity faces that could be extracted by averaging across identity trajectories (Fig. S4). Quantifying these slope changes (Fig. 2D, Fig. S5), we found that 147 (79.0%) of neurons showed significant slope difference between the caricature and anti-face sides. For each identity, we found a prevalence of V-shaped (21.8 %) and knee-shaped (10.8 %) response patterns (Fig. S6) and 119 (64.0%) of cells had at least one identity trajectory that showed V-shaped or knee-shaped pattern. Modeling with a combination of V-shaped and linear ramp tuning showed contribution of both components to the tuning shape of each identity trajectory (Fig. S7). Further, while individual cells varied in their species selectivity, the same population of cells exhibited similar V-shaped tuning for the human faces relative to the average human face (Fig, S8-10). As a control experiment, we morphed faces along trajectories between pairs of original identities and confirmed that neurons had no clear V-shaped tuning pattern if the identity trajectory did not pass through the average face

(Fig. S11). Critically, we showed, using data from the same experiment, as well data from a separate experiment exploiting long-term longitudinal recordings, that this tuning was present from the beginning of the session and thus could not be explained by within-session visual adaptation (*15, 16*) (Figs. 2E-G; see Supplementary Material).

We next turned our attention to the temporal evolution of this identity tuning profile. We tracked the time course of the V-shaped tuning using a sliding-window analysis (Fig. 3). Fig. 3A shows the time course of a typical face-selective neuron, showing a large transient response around 100 ms after stimulus onset followed by later sustained response. Note that for this neuron, the V-shaped profile was not clearly evident until late in the response, reaching a plateau only ~200 ms after stimulus presentation. This delayed profile was a shared feature across neurons and clearly evident in the population-averaged response, where V-shaped tuning emerged long after the initial transient response (Fig. 3B). Late tuning for facial identity is consistent with some previous studies reporting higher face identity tuning of neurons in inferotemporal cortex later in the neural response period (*20, 21*). It is also reminiscent of the more recent results in which the late responses of inferior temporal neurons are modulated by a hierarchical error signal (*17*). Although there were no differences in the initial response latency across identity level, the response magnitude to the low identity faces was abruptly attenuated during the visual response, while that to the high positive and negative identity faces persisted (Fig. 3C, D). Across the population, the number of V-shaped-tuned neurons began to increase around 100 ms and peaked at around 200 ms (55.8 ± 3.5% between 200 ± 50 ms, Fig. 3E), consistent with the shape of the population-averaged tuning. On average, V-shaped tuning emerged 54.8 ± 6.9 ms (mean ± SE) later than firing rate response (Fig. 3F).

We next asked whether the analysis of spike timing might lend insights into the mechanisms of delayed identity tuning. Our microwire bundles simultaneously recorded local populations of neurons all within a few hundred micrometers (*22*). We examined the spiking of neural pairs for their temporal coherence, which is thought to be an index of some aspects of neural coding, including underlying efficient information transmission from a population (*23-25*). Focusing first on the late visual response, we found that the spike-spike coherence between pairs of neurons (Fig. 4A) was notably stronger for low-identity faces than for more extreme faces (Fig. 4B, Fig. S12-14), suggesting a larger contribution from the local network during processing of low-identity faces. Coherent spikes were generally coincident in their timing, showing zero-phase coherence (Fig. 4C). Spike-LFP coherence was similarly heightened for low-identity faces (Fig. S13), with spikes most commonly issued in the troughs of the low-γ LFP when oscillatory inhibition should be weakest (*26*) (Fig.S13B). Both spike-spike and spike-LFP coherence were highest around the average face and lowest at higher identity levels, resulting in an inverted V-shaped pattern opposite to that of firing rate responses (Fig. 4D). Importantly, this elevated coherence emerged long after the initial visual response and grew coincident with the V-shaped tuning profile (Fig. 4E-F). These results could not have arisen from a trivial coupling of firing rate and coherence, as we matched the spiking rates in the coherence calculation (see Materials and Methods for details). We interpret the emergence of V-shaped tuning, and its coincidence with heightened spiking synchrony, as a manifestation of broad inhibition during the late response period (Figure 4G). Increases in spike timing synchrony among excitatory neurons have previously been linked to circuit inhibition (*26-30*). In this case, the suppression of the late response was strongest for faces with the lowest identity information. At a systems level, this

attenuation of face averageness serves to adjust or normalize the representation of identity to focus on features important for recognition.

Previous studies found a preponderance of linear or ramp tuning, but not tuning around the average face, among face patch neurons for features or principal components signaling identity (*14, 15*). One potential difference from the previous studies is our longer stimulus presentation time (300 ms), which may have given sufficient time for the delayed suppression (Figs. 3, 4) to emerge compared to the short presentations (117-150 ms) in previous studies. Here we found that the linear ramp tuning played a primary role during early transient response, while both the V-shaped tuning and linear ramp tuning played equally important role during later sustained response (Fig. S7F). It is unlikely that the species of the face stimulus is the basis of the difference, since we found qualitatively similar effects for both macaque and human faces, and since our previous findings used human faces (*10*). The method of analysis may also contribute to the observed differences. In the present study, we often summarized individual neuron data by averaging across multiple identity trajectories (Fig. 2, S2). This averaging together of V-shaped trajectories, "knee-shaped" trajectories, and linear ramp-shaped trajectories typically revealed a minimum response at the average face, even though some of the trajectories appeared with no discontinuity there (quantified in Figs. S4, S6 and S7).

We observed the characteristic V-shaped tuning across multiple face patches (Fig. 2C), suggesting that the identity normalization is a common feature in these regions. Although there was no clear difference in terms of the V-shaped tuning, some other feature of visual responses like tuning width was different between face patches (kurtosis values of 3.60 in AF vs. 4.81 in AM, measured across monkey face stimuli, $p<0.001$, Wilcoxon rank-sum test), as previous study reporting higher identity selectivity in AM face patch (*21*).

Normalization is an operation found in multiple areas of the brain including early visual cortices (*18*), where a range of neuronal responses is dynamically adjusted through the normalization process to maximize neural sensitivity. Here, we found a similar normalizing operation of face identity in higher visual areas of the inferotemporal cortex. Consistent with the psychological concept of norm-based face coding (*5-9*), the identity normalization can increase neuronal sensitivity to faces in a familiar group of people (*9*). While early visual cortices utilize visual input itself for the normalization of the visual input (*18*), the identity normalization relies on subjects' past experiences which works as a top-down modulation signal, similarly to normalization of attention through a top-down signal (*31*).

How the brain gradually gathers information about important stimuli such as faces to generate norms for certain visual problems is currently a matter of conjecture. For faces, one possibility is that this information is learned and stored external to the visual system proper, perhaps in an area such as the perirhinal cortex that shares dense reciprocal connections with the inferior temporal cortex (*32*). Recent work showing that a face-specific subregion of the perirhinal cortex responds selectively to familiar faces (*33*) suggests that this area might provide an external feedback signal to face patches, perhaps normalizing responses according to face averageness. Another possibility is that the disparate face patches exhibit a parallel learning of facial structure, mainly through experience during development, which they then apply locally through local inhibitory mechanisms. Future work may address whether this mode of information encoding is unique to faces, or whether it may also pertain to other stimuli such as color, shape, or bodily postures (*34-36*). Likewise, an intriguing possibility for future study is that this mechanism for face coding in primates reflects a more general principle of how brains apply gradually learned

statistics of important stimulus categories, and particularly those related to the perception of conspecifics, an aspect of visual cognition thought to be critical for sexual selection (*1, 37, 38*).

References and Notes:

1. D. A. Leopold, G. Rhodes, A comparative view of face perception. *J Comp Psychol* **124**, 233-251 (2010).
2. N. Kanwisher, J. McDermott, M. M. Chun, The fusiform face area: a module in human extrastriate cortex specialized for face perception. *J Neurosci* **17**, 4302-4311 (1997).
3. D. Y. Tsao, W. A. Freiwald, T. A. Knutsen, J. B. Mandeville, R. B. Tootell, Faces and objects in macaque cerebral cortex. *Nat Neurosci* **6**, 989-995 (2003).
4. C. C. Hung *et al*., Functional mapping of face-selective regions in the extrastriate visual cortex of the marmoset. *J Neurosci* **35**, 1160-1172 (2015).
5. P. J. Benson, D. I. Perrett, Perception and recognition of photographic quality facial caricatures: Implications for the recognition of natural images. *European Journal of Cognitive Psychology* **3**, 105-135 (1991).
6. G. Rhodes, D. A. Leopold, in *The Oxford handbook of face perception,* A. Calder, G. Rhodes, J. M., H. J., Eds. (Oxford University Press, Oxford, 2011), pp. 263-286.
7. D. A. Leopold, A. J. O'Toole, T. Vetter, V. Blanz, Prototype-referenced shape encoding revealed by high-level aftereffects. *Nat Neurosci* **4**, 89-94 (2001).
8. M. A. Webster, D. Kaping, Y. Mizokami, P. Duhamel, Adaptation to natural facial categories. *Nature* **428**, 557-561 (2004).
9. T. Valentine, A unified account of the effects of distinctiveness, inversion, and race in face recognition. *Q J Exp Psychol A* **43**, 161-204 (1991).
10. D. A. Leopold, I. V. Bondar, M. A. Giese, Norm-based face encoding by single neurons in the monkey inferotemporal cortex. *Nature* **442**, 572-575 (2006).
11. G. Loffler, G. Yourganov, F. Wilkinson, H. R. Wilson, fMRI evidence for the neural representation of faces. *Nat Neurosci* **8**, 1386-1390 (2005).
12. M. Latinus, P. McAleer, P. E. Bestelmeyer, P. Belin, Norm-based coding of voice identity in human auditory cortex. *Curr Biol* **23**, 1075-1080 (2013).
13. S. Panis, J. Wagemans, H. P. Op de Beeck, Dynamic norm-based encoding for unfamiliar shapes in human visual cortex. *J Cogn Neurosci* **23**, 1829-1843 (2011).
14. L. Chang, D. Y. Tsao, The Code for Facial Identity in the Primate Brain. *Cell* **169**, 1013-1028 e1014 (2017).
15. W. A. Freiwald, D. Y. Tsao, M. S. Livingstone, A face feature space in the macaque temporal lobe. *Nat Neurosci* **12**, 1187-1196 (2009).
16. D. A. Kahn, G. K. Aguirre, Confounding of norm-based and adaptation effects in brain responses. *Neuroimage* **60**, 2294-2299 (2012).
17. E. B. Issa, C. F. Cadieu, J. J. DiCarlo, Neural dynamics at successive stages of the ventral visual stream are consistent with hierarchical error signals. *Elife* **7**, (2018).
18. M. Carandini, D. J. Heeger, Normalization as a canonical neural computation. *Nat Rev Neurosci* **13**, 51-62 (2011).

19. V. Blanz, A. J. O'Toole, T. Vetter, H. A. Wild, On the other side of the mean: the perception of dissimilarity in human faces. *Perception* **29**, 885-891 (2000).
20. Y. Sugase, S. Yamane, S. Ueno, K. Kawano, Global and fine information coded by single neurons in the temporal visual cortex. *Nature* **400**, 869-873 (1999).
21. W. A. Freiwald, D. Y. Tsao, Functional compartmentalization and viewpoint generalization within the macaque face-processing system. *Science* **330**, 845-851 (2010).
22. D. B. McMahon, I. V. Bondar, O. A. Afuwape, D. C. Ide, D. A. Leopold, One month in the life of a neuron: longitudinal single-unit electrophysiology in the monkey visual system. *J Neurophysiol* **112**, 1748-1762 (2014).
23. K. W. Koyano *et al*., Laminar Module Cascade from Layer 5 to 6 Implementing Cue-to-Target Conversion for Object Memory Retrieval in the Primate Temporal Cortex. *Neuron* **92**, 518-529 (2016).
24. G. G. Gregoriou, S. J. Gotts, H. Zhou, R. Desimone, High-frequency, long-range coupling between prefrontal and visual cortex during attention. *Science* **324**, 1207-1210 (2009).
25. H. L. Dean, M. A. Hagan, B. Pesaran, Only coherent spiking in posterior parietal cortex coordinates looking and reaching. *Neuron* **73**, 829-841 (2012).
26. V. S. Sohal, How Close Are We to Understanding What (if Anything) gamma Oscillations Do in Cortical Circuits? *J Neurosci* **36**, 10489-10495 (2016).
27. M. A. Whittington, R. D. Traub, N. Kopell, B. Ermentrout, E. H. Buhl, Inhibition-based rhythms: experimental and mathematical observations on network dynamics. *Int J Psychophysiol* **38**, 315-336 (2000).
28. A. Hasenstaub, S. Otte, E. Callaway, Cell Type-Specific Control of Spike Timing by Gamma-Band Oscillatory Inhibition. *Cereb Cortex* **26**, 797-806 (2016).
29. A. Hasenstaub *et al*., Inhibitory postsynaptic potentials carry synchronized frequency information in active cortical networks. *Neuron* **47**, 423-435 (2005).
30. V. S. Sohal, F. Zhang, O. Yizhar, K. Deisseroth, Parvalbumin neurons and gamma rhythms enhance cortical circuit performance. *Nature* **459**, 698-702 (2009).
31. J. H. Reynolds, D. J. Heeger, The Normalization Model of Attention. *Neuron* **61**, 168-185 (2009).
32. P. Lavenex, W. A. Suzuki, D. G. Amaral, Perirhinal and parahippocampal cortices of the macaque monkey: projections to the neocortex. *J Comp Neurol* **447**, 394-420 (2002).
33. S. M. Landi, W. A. Freiwald, Two areas for familiar face recognition in the primate brain. *Science* **357**, 591-595 (2017).
34. A. H. Bell, F. Hadj-Bouziane, J. B. Frihauf, R. B. Tootell, L. G. Ungerleider, Object representations in the temporal cortex of monkeys and humans as revealed by functional magnetic resonance imaging. *J Neurophysiol* **101**, 688-700 (2009).
35. K. Srihasam, J. B. Mandeville, I. A. Morocz, K. J. Sullivan, M. S. Livingstone, Behavioral and anatomical consequences of early versus late symbol training in macaques. *Neuron* **73**, 608-619 (2012).
36. R. Lafer-Sousa, B. R. Conway, Parallel, multi-stage processing of colors, faces and shapes in macaque inferior temporal cortex. *Nat Neurosci* **16**, 1870-1878 (2013).
37. B. J. Bradley, N. I. Mundy, The primate palette: The evolution of primate coloration. *Evol Anthropol* **17**, 97-111 (2008).
38. G. Rhodes, The evolutionary psychology of facial beauty. *Annu Rev Psychol* **57**, 199-226 (2006).

39. D. B. McMahon, A. P. Jones, I. V. Bondar, D. A. Leopold, Face-selective neurons maintain consistent visual responses across months. *Proc Natl Acad Sci U S A* **111**, 8251-8256 (2014).
40. B. E. Russ, D. A. Leopold, Functional MRI mapping of dynamic visual features during natural viewing in the macaque. *Neuroimage* **109**, 84-94 (2015).
41. W. F. Asaad, N. Santhanam, S. McClellan, D. J. Freedman, High-performance execution of psychophysical tasks with complex visual stimuli in MATLAB. *J Neurophysiol* **109**, 249-260 (2013).
42. J. Hwang, A. R. Mitz, E. A. Murray, NIMH MonkeyLogic: Behavioral control and data acquisition in MATLAB. *J Neurosci Methods* **323**, 13-21 (2019).
43. M. Kleiner, D. Brainard, D. Pelli, What's new in Psychtoolbox-3? *Perception* **36**, 14-14 (2007).
44. I. V. Bondar, D. A. Leopold, B. J. Richmond, J. D. Victor, N. K. Logothetis, Long-term stability of visual pattern selective responses of monkey temporal lobe neurons. *PLoS One* **4**, e8222 (2009).
45. P. Mitra, H. Bokil, *Observed brain dynamics*. (Oxford University Press, New York ; Oxford, 2008), pp. xxii, 381 p.
46. P. P. Mitra, B. Pesaran, Analysis of dynamic brain imaging data. *Biophys J* **76**, 691-708 (1999).
47. P. Berens, CircStat: A MATLAB Toolbox for Circular Statistics. *J Stat Softw* **31**, 1-21 (2009).
48. J. F. Mitchell, K. A. Sundberg, J. H. Reynolds, Spatial attention decorrelates intrinsic activity fluctuations in macaque area V4. *Neuron* **63**, 879-888 (2009).

Acknowledgments:

We would like to thank David Yu, Charles Zhu and Frank Ye for assistance with fMRI scanning, Julie Hong, Katy Smith, George Dold and David Ide for technical assistance and development, Olga Dal Monte and Carlos E. Thomaz for providing face stimli, Vasileva Liubov, Igor Bondar, David Godlove, Bruno Averbeck, Teppei Matsui and Toshiyuki Hirabayashi for helpful discussions. Functional and anatomical MRI scanning was carried out in the Neurophysiology Imaging Facility Core (NIMH, NINDS, NEI). This work utilized the computational resources of the NIH HPC Biowulf cluster http://hpc.nih.gov). **Funding:** This work was supported by the Intramural Research Program of the National Institute of Mental Health (ZIAMH002838 and ZIAMH002898) to D.A.L. K.W.K. was supported by the fellowship from the Uehara Memorial Foundation. **Author contributions:** A.P.J. and D.A.L. designed the experiments. A.P.J. prepared

stimulus set. K.W.K., D.B.T.M., E.N.W., B.E.R. and D.A.L. performed animal surgeries. K.W.K., D.B.T.M. and B.E.R. run fMRI experiments and analysis. K.W.K, A.P.J. and E.N.W. collected electrophysiology data. K.W.K, A.P.J. and D.A.L. performed analysis of neuronal data. K.W.K and D.A.L. wrote the paper. All the authors contributed to editing of the manuscript. D.A.L. supervised the study. **Competing interests:** Authors declare no competing interests. **Data and materials availability:** All the data in the paper are presented in the main text or supplementary materials.

Supplementary Materials:
Materials and Methods
Figures S1-S15
References (39-48)

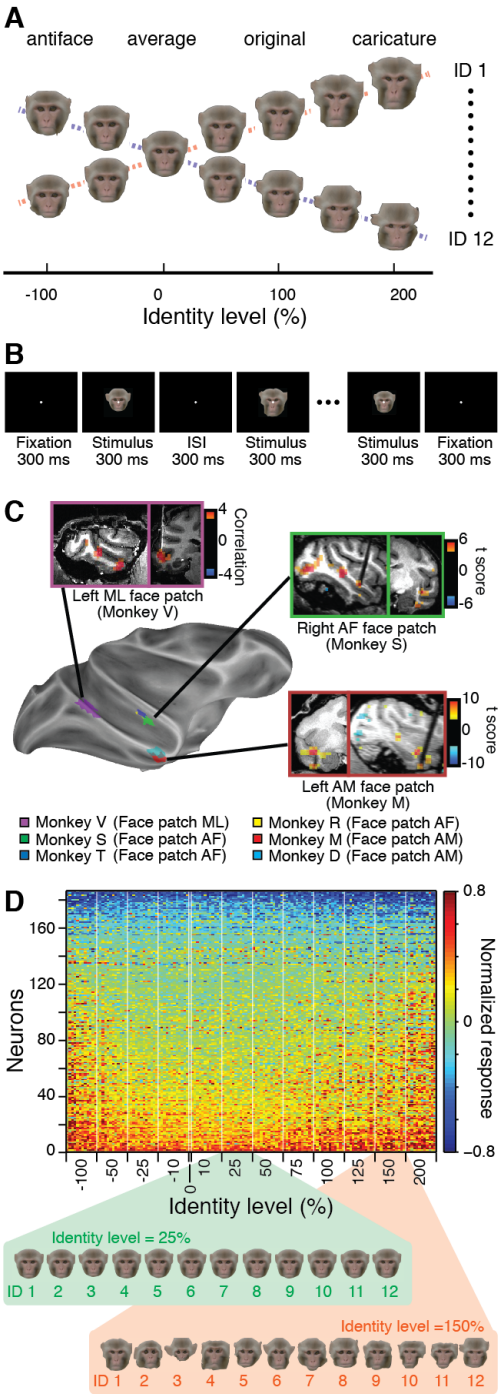

Figure 1. Experimental design. (**A**) Stimulus set. An average face (0% identity level) was first created by averaging 12 faces. Then faces were morphed between the average and each of the 12 original (100%) identities, creating caricatures (200%) and antifaces (-100%) which have exaggerated and opposite facial features of the original, respectively. (**B**) Sequence of passive-viewing task. (**C**) Face patches from where neurons were recorded. (**D**) Responses of face-selective neurons (n=186) for monkey face stimuli. This raw depiction of the data clearly shows that neurons exhibit smaller responses to low-identity-level faces.

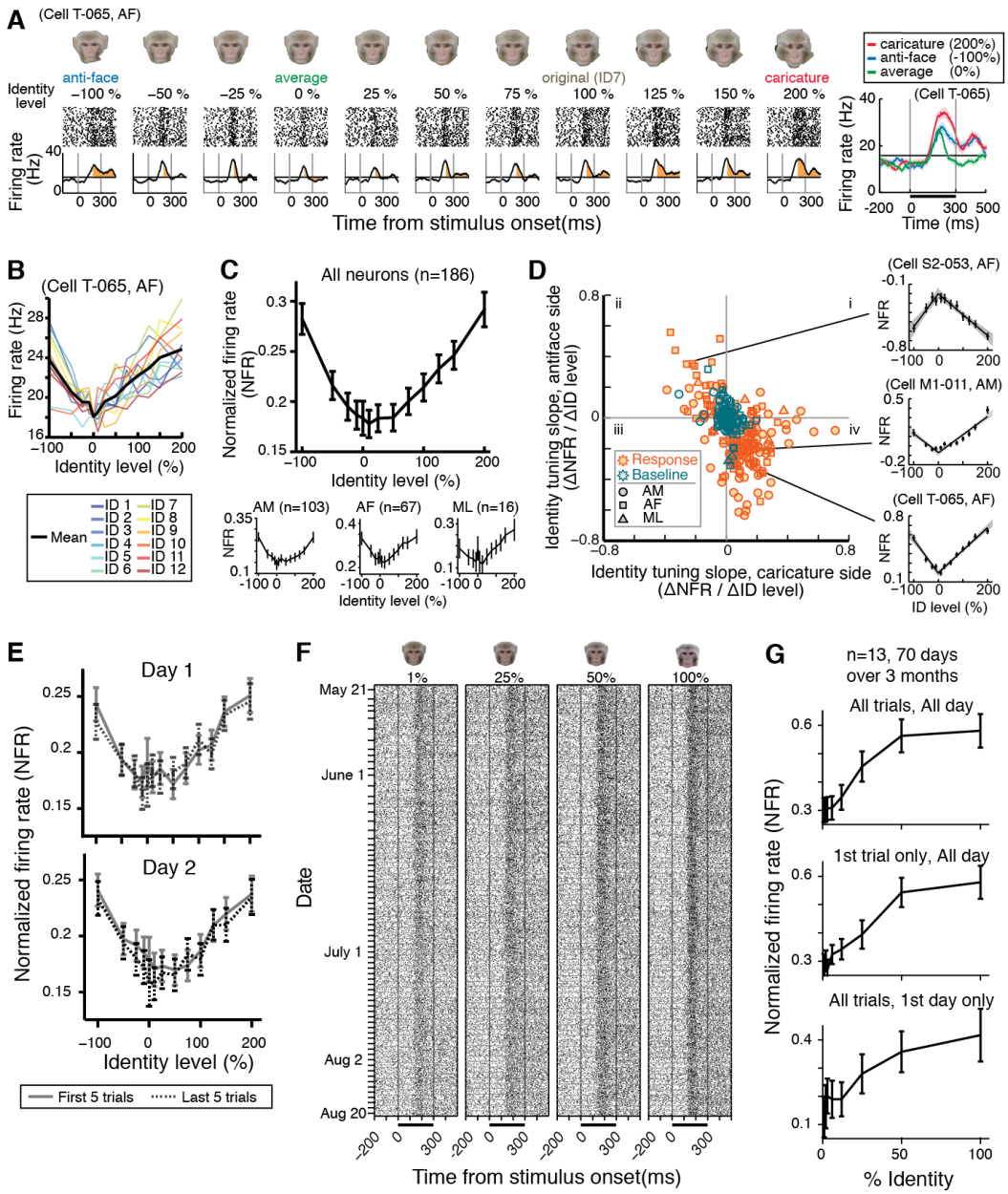

Figure 2. V-shaped tuning by face patch neurons. (**A**) Response of a neuron in an identity trajectory, exhibiting smallest response to the average face. (**B**) Response of the neuron in (A) for all the identity trajectories. The mean response showed "V-shaped" tuning pattern along the identity level. (**C**) Mean response of all the face-selective neurons, showing characteristic V-shaped tuning. Insets, mean of each of AM, AF and ML patches. (**D**) Quantification of tuning shape by two linear regression lines. During response period, slopes in caricature side tended to be anti-correlated with slopes in antiface side, indicating V-shaped or Λ–shaped tuning. (**E**) V-shaped tuning pattern consistently observed through recording sessions. (**F**) Longitudinal recording from a neuron over 3 months, showing stable tuning pattern along identity levels. (**G**) Tuning pattern consistently observed through the longitudinal recording session. The response was smallest to the average face even in the first trial of each day and first day of the recording, which cannot be explained by adaptation or familiarity effects.

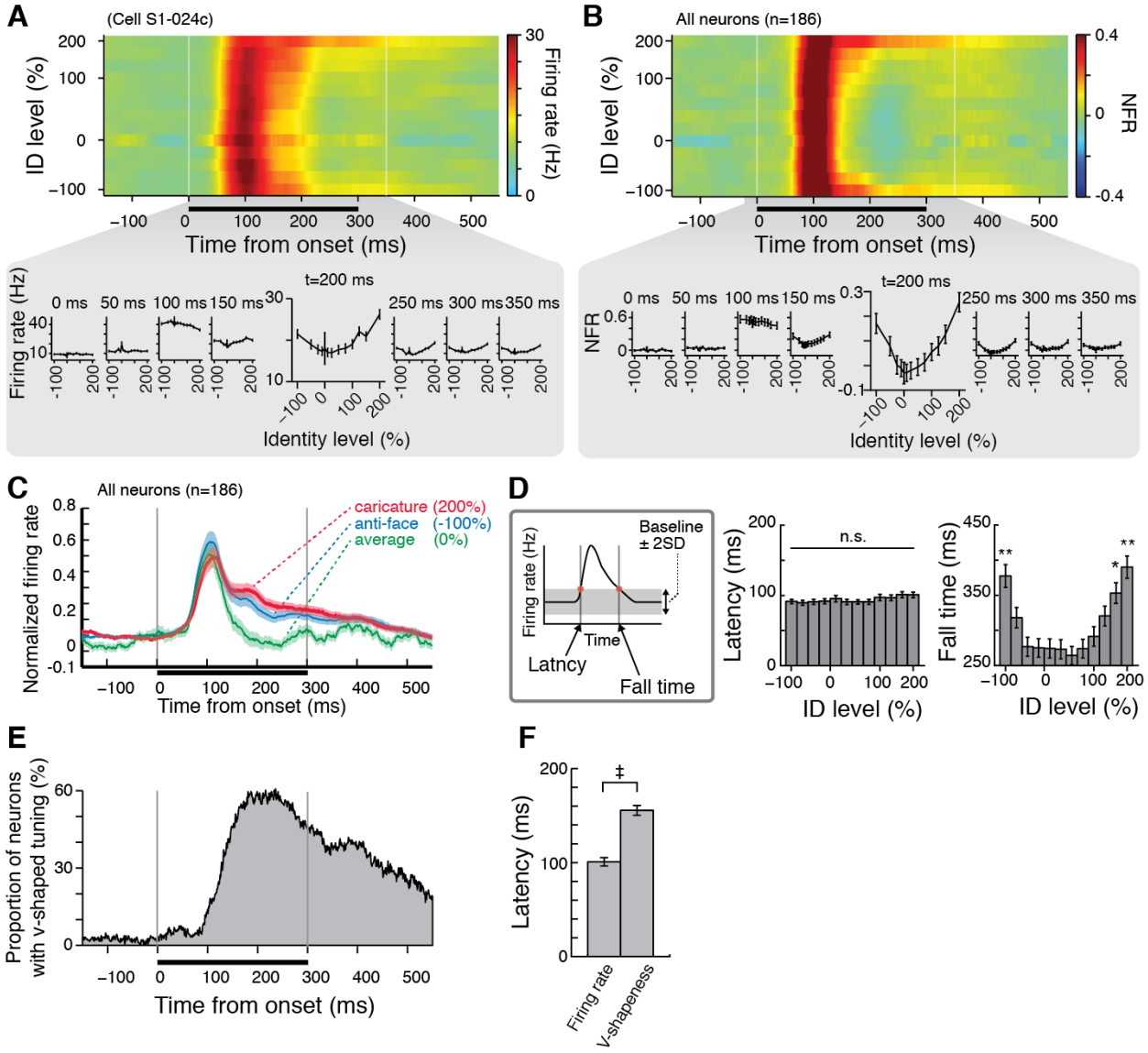

Figure 3. Late emergence of V-shaped tuning. (**A** and **B**) timecourse of the v-shaped tuning pattern in an example neuron (A) and population average (B). The V-shaped tuning pattern did not clearly observed during the large transient response around 100 ms but slowly emerged during later sustained period that started around 150 ms. (**C**) Mean response timecourses of average, caricatures and antifaces, which diverges in the sustained response period. (**D**) Response latency and fall-time. The response dropped earlier for faces of lower identity levels. **, p<0.01 (Tukey's post-hoc test after one-way ANOVA) (**E**) Number of V-shaped tuning cells that increased at later sustained period. (**F**) Latency of V-shapeness, which emerges later than firing rate. ‡, p<0.01 (paired t-test with Bonferroni's correction).

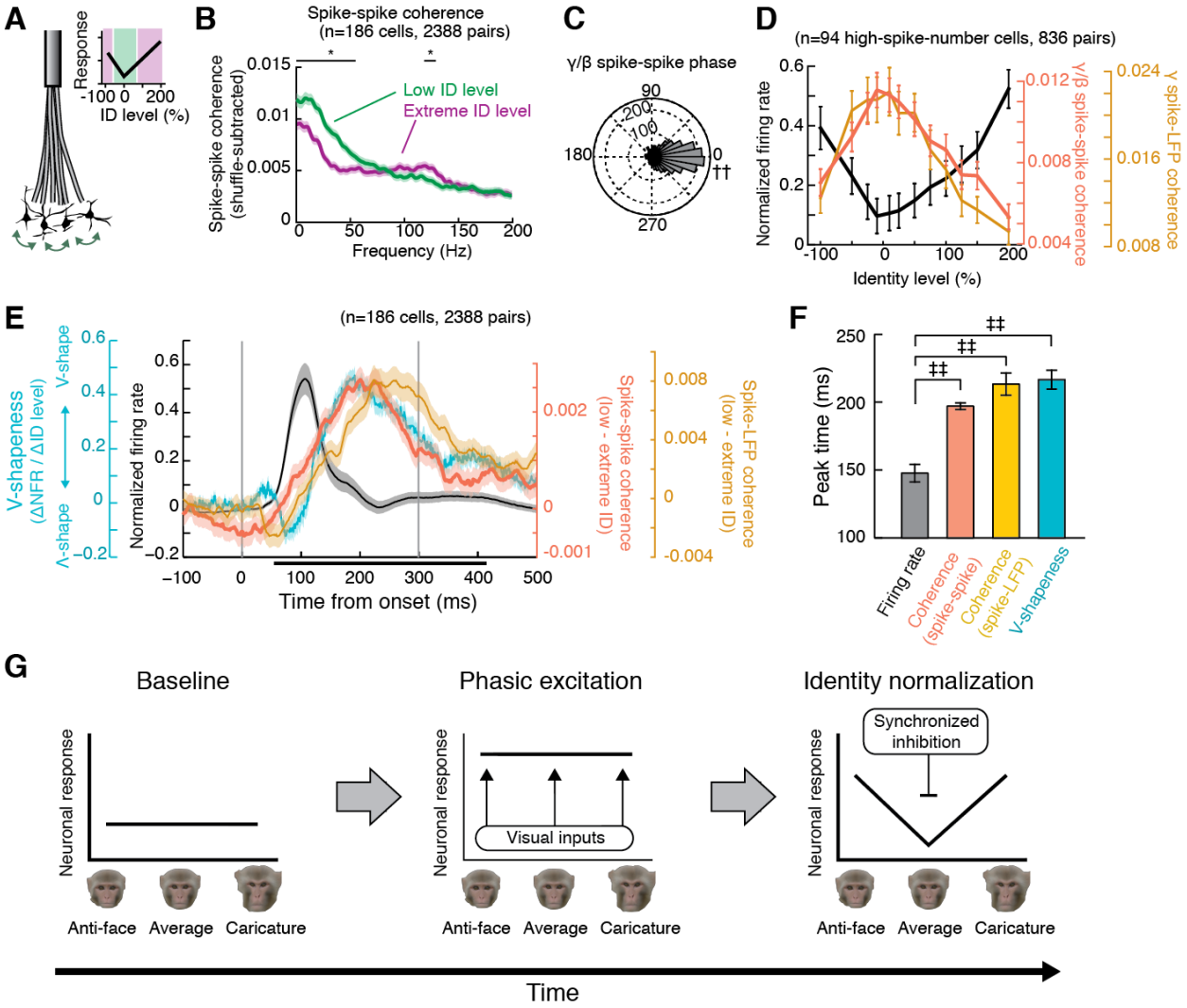

Figure 4. Neuronal synchrony for the average face. (**A**) Simultaneous recording of face patch neurons. Inset, classification of low and extreme identity levels. (**B**) Frequency spectrum of spike-spike coherence between simultaneously recorded neuron pairs. The coherence was higher in low identity level than extreme identity level at <50 Hz. **, p<0.05 (paired t-test with Bonferroni's correction). (**C**) Phase relationships between neuronal pairs, showing coherent firing of spikes. ††, p<0.01 (Rayleigh test). (**D**) Highest coherence for the average face, showing opposite tuning pattern to firing rate. (**E**) Coherence timecourse following v-shapeness timecourse rather than firing rate timecourse. (**F**) Peak time of coherence which is longer than firing rate. ‡‡, p<0.01 (paired t-test with Bonferroni's correction). (**G**), Model of neuronal mechanism for identity normalization, shaped by general phasic excitation followed by synchronized inhibition for the average face.