

# How Do We Recognize a Face?

Rodrigo Quian Quiroga<sup>1,\*</sup>

<sup>1</sup>Centre for Systems Neuroscience, University of Leicester, 9 Salisbury Rd., Leicester LE1 7QR, UK

\*Correspondence: [rqqg1@le.ac.uk](mailto:rqqg1@le.ac.uk)

<http://dx.doi.org/10.1016/j.cell.2017.05.012>

**How individual faces are encoded by neurons in high-level visual areas has been a subject of active debate. An influential model is that neurons encode specific faces. However, Chang and Tsao conclusively show that, instead, these neurons encode features along specific axes, which explains why they were previously found to respond to apparently different faces.**

We arrive at a lab reunion dinner in a well-attended conference, and very rapidly and without effort, we recognize the faces of our previous lab-mates, in spite of the fact that we see some from the front, others from profile, and that some have changed their appearance since we last saw them. Contrasting with how natural it seems to perform such a task, face recognition involves complex processes (Logothetis and Sheinberg, 1996) that we still do not fully understand. Converging evidence has shown that visual perception involves the processing of information along the “ventral visual pathway,” ending at the infero-temporal cortex (IT), where neurons were shown to respond to body parts and faces (Gross et al., 1972). Within this area, a major breakthrough was the finding of anatomically segregated patches with neurons that respond nearly exclusively to faces (Tsao et al., 2006). But how do neurons in these patches encode the people we recognize? A plausible idea, known as the “exemplar-based model,” is that neurons encode specific individuals—i.e., a neuron responds preferentially to a particular face and with lesser strength the more the presented face differs from the preferred one (Valentine, 1991). In line with this view, neurons in face patch AM (anterior medial)—the final stage of face processing in the IT—were shown to respond selectively to specific faces (Freiwald and Tsao, 2010). Such coding brings to mind the notion of “grandmother cells” (Gross, 2002), by which processing along the ventral visual pathway leads to increasingly complex representations that would end up with neurons responding only to a specific person, such as our grandmother (no matter how we see her). However, a puzzling result was

the fact that neurons in AM commonly showed responses to more than one face, without any apparent similarity between these. The article by Chang and Tsao in this issue of *Cell* (Chang and Tsao, 2017) solves this puzzle by demonstrating the existence of a very elegant and straightforward coding mechanism by neurons in high-level visual areas.

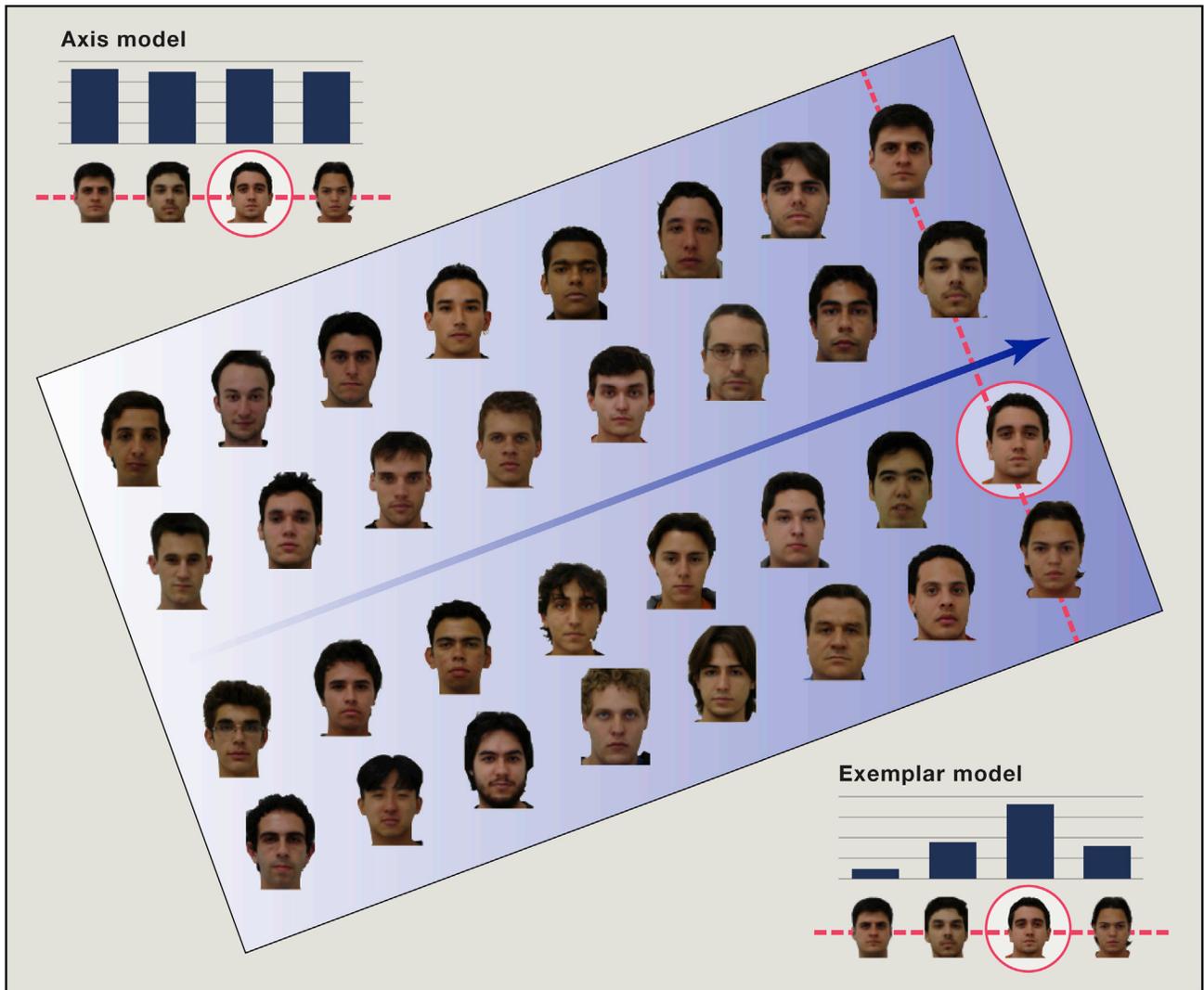
Experiments involved single-neuron recordings of face-responsive neurons in patches AM, ML (middle lateral), and MF (middle fundus)—identified from previous fMRI localizer sessions (Tsao et al., 2006)—while monkeys were shown 2,000 faces. With an impressive set of results and analyses, Chang and Tsao (2017) conclusively show that, rather than an exemplar-based coding, neurons in these face patches show an “axis coding,” firing according to the projection of the presented faces onto axes representing specific features. Moreover, face-patch AM carried complementary information compared to patches ML and MF: AM gave information about the appearance of the faces, whereas ML and MF provided information about their shape.

The model by Chang and Tsao (2017) leads naturally to a very interesting prediction: in a direction orthogonal to the feature-axis coded by a neuron, the neuron’s tuning should be flat; all that matters is the projection along the feature-axis. Figure 1 illustrates the idea. Let us assume that a given neuron encodes a feature represented by the blue axis and that a response is found to the circled face. Then, following the model of Chang and Tsao (2017), this neuron should respond with similar strength to the other faces along the orthogonal axis (blue dotted line), which is exactly what was found. On the contrary, according to the

exemplar-based model, there should be a non-flat tuning for the faces along this dimension, which was not the case.

Given that different neurons encode different features, from a population of neurons, it should then be possible to distinguish different faces. Taking the activity of a population of about 200 face cells, Chang and Tsao (2017) showed that they could not only decode but also reproduce the presented faces with impressive accuracy. Moreover, tackling the long-disputed issue of visual invariance—i.e., the neural mechanisms that allow us to recognize the same person under varying viewing conditions (Logothetis and Sheinberg, 1996)—they show that they could map features of front and profile views and then decode different views of the same individuals based on the neurons’ activity.

Chang and Tsao (2017) explored another interesting prediction of their model. Once knowing the feature that is encoded by a given neuron, it should be possible to predict the response to faces not previously shown. This is what they found; all they needed to do was to calculate the projection of the faces onto the corresponding feature axes. Altogether, these findings have general implications about estimations of neuronal selectivity in high-level visual areas. Neurons in face patch AM are the ones with the sharper tuning to faces in the visual system (Freiwald and Tsao, 2010), but what the new results show is that estimations of neural selectivity can in principle give any outcome, depending on the projections onto the feature axis of the faces of the stimulus set used. As the authors argue, their results imply that there are no detectors for face identity at the single-neuron level in the face patch system



**Figure 1. Illustration of the Face Recognition “Axis Model” by Chang and Tsao**

Consider an axis (in blue) representing the visual feature encoded by a neuron and a set of faces in this space. (Faces were obtained from the FEI face database and were arbitrarily placed at regular intervals for simplicity.) Following the “axis model,” the neuron should have a similar response to the circled face and the ones along the same orthogonal axis (blue dotted line). In contrast, if neurons respond preferentially to specific faces (the “exemplar model”), then the tuning of the neuron along the orthogonal axis should not be flat. Chang and Tsao conclusively show that the former is the case.

and, consequently, this may put an end to the long-standing dispute about the existence of grandmother cells in visual cortex.

A final point of interest is the fact that neurons in the face patches are just a couple of synapses away from the medial temporal lobe (MTL; the hippocampus and surrounding cortex), where, in humans, neurons were found to encode concepts, responding in a very selective and abstract way to specific persons or places (Quian Quiroga et al., 2005). More recently, it was also shown that these neurons respond to related concepts,

thus encoding meaningful associations in a high-dimensional cognitive space (De Falco et al., 2016). The encoding of such associations has been proposed to be the basis of declarative memory functions (Quian Quiroga, 2012), and compared to the coding showed by Chang and Tsao (2017), one can pose a fascinating question: how does the brain go from a perception-driven representation of features in the face patches to a memory-driven representation of concepts and their associations in the MTL? One possibility is that there is a fundamental difference between species or,

alternatively, that such difference is given by the fact that, for humans, the faces of people we know (to which MTL neurons tend to respond) convey meaning with associated memories and emotions, compared to the “meaningless” faces that were used for these experiments with monkeys, with the former type of stimuli giving a representation (like the exemplar-based model) closer to the one found in the human MTL. Another more thrilling possibility is that the exquisite structure of the hippocampal formation—and particularly, the recurrent connectivity in area CA3—gives rise to a

major change in the metric space and that such change provides in turn the substrate of two different, though intrinsically related, functions: perception and memory.

In sum, the article of [Chang and Tsao \(2017\)](#) not only provides a major breakthrough in our understanding of how faces are encoded by neurons in high-level visual areas, but also poses new, fascinating questions and challenges that are ripe for future research.

## REFERENCES

- Chang, L., and Tsao, D. (2017). *Cell* 169, this issue, 1013–1028.
- De Falco, E., Ison, M.J., Fried, I., and Quiari Quiruga, R. (2016). *Nat. Commun.* 7, 13408.
- Freiwald, W.A., and Tsao, D.Y. (2010). *Science* 330, 845–851.
- Gross, C.G. (2002). *Neuroscientist* 8, 512–518.
- Gross, C.G., Rocha-Miranda, C.E., and Bender, D.B. (1972). *J. Neurophysiol.* 35, 96–111.
- Logothetis, N.K., and Sheinberg, D.L. (1996). *Annu. Rev. Neurosci.* 19, 577–621.
- Quiari Quiroga, R. (2012). *Nat. Rev. Neurosci.* 13, 587–597.
- Quiari Quiroga, R., Reddy, L., Kreiman, G., Koch, C., and Fried, I. (2005). *Nature* 435, 1102–1107.
- Tsao, D.Y., Freiwald, W.A., Tootell, R.B., and Livingstone, M.S. (2006). *Science* 311, 670–674.
- Valentine, T. (1991). *Q. J. Exp. Psychol. A.* 43, 161–204.

# Noninvasive Neuromodulation Goes Deep

Jacek Dmochowski<sup>1</sup> and Marom Bikson<sup>1,\*</sup>

<sup>1</sup>Department of Biomedical Engineering, The City College of New York of CUNY, New York, NY 10031, USA

\*Correspondence: [bikson@ccny.cuny.edu](mailto:bikson@ccny.cuny.edu)

<http://dx.doi.org/10.1016/j.cell.2017.05.017>

**Modulating deep regions of the brain with noninvasive technology has challenged researchers for decades. In a new study, Grossman et al. leverage the emergence of a slowly oscillating “beat” from intersecting high-frequency electric fields to stimulate deep brain regions, opening a frontier in the biophysics and technology of brain stimulation.**

As tools to treat neuropsychiatric disorders and to probe cognition, noninvasive techniques for stimulating the brain are an area of intense development. Electromagnetic approaches such as transcranial magnetic stimulation and transcranial direct current stimulation are at the forefront. One seemingly unavoidable limitation of these and other electromagnetic techniques is the inability to target deep brain regions. With noninvasive stimulation, the intensity of electromagnetic fields drops off with distance from the surface of the head, meaning that superficial brain areas are activated first. In this issue of *Cell*, [Grossman et al. \(2017\)](#) exploit a long-standing acoustical phenomenon to propose a form of noninvasive electrical brain stimulation capable of stimulating deep brain areas in a selective manner: “temporal interference” (TI) stimulation.

When two tones of similar frequency are simultaneously emitted, the envelope of the net signal oscillates at a low frequency equal to the difference of the two tones, known as the “beat frequency.” The au-

thors adapt this concept to electrical stimulation, using two pairs of surface electrodes to concurrently apply 2 kHz and 2.01 kHz sinusoidal stimulation. This creates three classes of brain regions—two superficial regions exposed to an electric field at either 2 or 2.01 kHz with a fixed amplitude (flat envelope) and an intermediate brain region with a net electric field whose envelope oscillates at 10 Hz, the beat frequency ([Figure 1](#)). The underlying assumption of TI stimulation is that neurons are unresponsive to high-frequency stimulation with a flat envelope, while neurons will respond to the “beating” high-frequency stimulation, leading to selective stimulation of neurons in the intermediate (i.e., deep) brain region.

Validation is not trivial: the same principle that renders high frequency ineffective, namely the low-pass properties of neuronal membranes ([Bikson et al., 2004](#); [Deans et al., 2007](#)), would a priori be predicted to equally attenuate a beating high-frequency waveform.

[Grossman et al. \(2017\)](#) first applied TI stimulation to the skulls of anesthetized mice, reporting that stimulation with 2 and 2.01 kHz evokes firing at a rate of 10 Hz (matching the difference of the two applied frequencies) in the somatosensory cortex and hippocampus, as measured by patch clamp electrophysiology. In separate experiments, the authors used *c-fos* labeling to show that TI stimulation preferentially activates hippocampal over cortical neurons. In each case, control experiments with only 2 kHz stimulation produced no neuronal response, while direct 10 Hz stimulation produced diffuse activation. Supporting safety, there was no evidence for pathological activity, brain temperature increase, or markers of brain injury.

The authors address the ability to “steerably” target various brain regions by changing electrode position or the applied current intensity. Finite element simulations of the electric field in a phantom and mouse head as well as in vivo experiments in mouse all suggest that spatial targeting