Medial Prefrontal Cortex Updates Its Status

Mina Cikara[1] and Samuel J. Gershman[1,*]
[1]Department of Psychology, Harvard University, Cambridge, MA 02138, USA
*Correspondence: gershman@fas.harvard.edu
http://dx.doi.org/10.1016/j.neuron.2016.11.040

How does the brain infer social status? A new study by Kumaran et al. (2016) identifies a region of the medial prefrontal cortex that, in concert with the amygdala and hippocampus, subserves updating of probabilistic beliefs about the status of individuals in a social hierarchy.

The ability to infer status—position in a social hierarchy—is paramount to survival: people who occupy higher positions in a hierarchy have greater access to resources than those beneath them, and therefore pose a more credible threat to other hierarchy members (Fiske, 1992). Status inference also allows individuals and groups to avoid engaging with foes to whom they are sure to lose. Reflecting the importance of this ability, even preverbal infants are sensitive to relative dominance cues (e.g., size) when two agents are in conflict (Thomsen et al., 2011).

How are these hierarchies learned in the first place? And how do these representations change depending on whether they are self-relevant or not? Finally, once they're learned, how readily are status representations retrieved? In this issue, Kumaran et al. (2016) provide a computational perspective on these questions, suggesting that the medial prefrontal cortex (mPFC), in coordination with the amygdala and hippocampus, serves as a status inference engine.

Kumaran et al. trained participants via trial and error to discriminate the relative status of individuals within two different social hierarchies: one to which participants belonged (the self-hierarchy) and one to which a close friend belonged (the other hierarchy). During training, the individuals always occupied adjacent positions in the hierarchy, but participants were intermittently tested (without feedback) on discriminations between individuals occupying non-adjacent positions (e.g., participants should identify the sixth ranked individual as lower in status than the third ranked). Accurate performance on these test trials demonstrated that participants acquired a representation of the hierarchy that enabled transitive inferences about status.

These results could be accounted for by a Bayesian model in which probabilistic beliefs about the hierarchy were updated based on the relative ranks observed on each training trial (Figure 1). This model embodies two important assumptions about status inference: (1) humans represent uncertainty about status, and (2) information about the relative status of two individuals is informative about the distribution of all individuals in the hierarchy. To account for gradual learning (or, equivalently, gradual forgetting), the authors posited a generative model in which status drifts slowly over time. As in other Bayesian models, Kumaran et al.'s model allowed uncertainty to impinge on behavior by governing both the learning rate (Behrens et al., 2007) and the slope of the decision function (De Martino et al., 2013). Accordingly, this model offered a quantitatively better account of the behavioral data compared to a reinforcement learning model that tracked point estimates of status. This was especially evident early in training, when uncertainty (represented by the Bayesian model, but not by the reinforcement learning model) was greatest.

The Bayesian model was used to generate trial-by-trial regressors for fMRI activity collected while participants performed the training and testing tasks. During training trials, a hierarchy update regressor (measuring how much the probabilistic belief changed as a consequence of feedback) correlated with activity in the mPFC, amygdala, and hippocampus. Closer inspection of the data revealed that the mPFC correlation was driven specifically by the self-hierarchy condition. Furthermore, the mPFC was coupled to a greater extent with the amygdala and hippocampus in the self-hierarchy relative to the other hierarchy condition. During test trials, the difference in inferred status of individuals correlated with activity in the amygdala and hippocampus. These same regions correlated with inferred status during a subsequent categorization task in which participants judged whether a face belonged to the self- or other hierarchy.

These results bring us closer to answering the questions posed above. First, the Bayesian model for status inference provides a behaviorally validated account of how social hierarchies can be learned from observations of one's own and others' relative status. The model's computations appear to depend on a network centered on the amygdala, hippocampus, and mPFC. Interestingly, the hippocampus has previously been implicated in representations of targets' social distance from oneself (e.g., whether the target is a close friend versus a distant acquaintance). For example, Tavares et al. (2015) asked participants to play a choose-your-own-adventure-type game in which participants interacted with a series of agents. In some interactions, participants decided whether to affiliate with an agent; in other interactions, they decided whether to comply with an agent's request. They discovered that the hippocampus tracked interaction-by-interaction updating of the social relationships between participants and each agent.

In a related study, Zink et al. (2008) observed bilateral occipital/parietal cortex, ventral striatum, parahippocampal cortex, and dorsolateral prefrontal cortex engagement when participants passively viewed higher- versus lower-ranked players in a game-based hierarchy. Converging with Kumaran et al.'s findings, they additionally

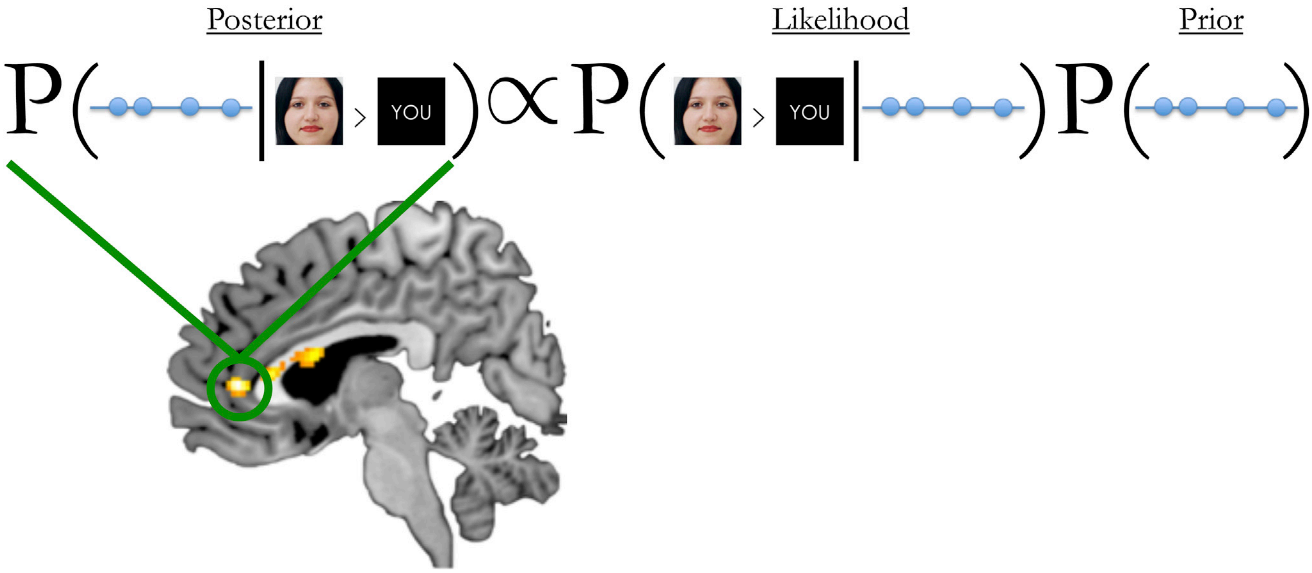

Figure 1. A Bayesian Model of Status Inference
Given information about the relative status of two individuals, the model updates its beliefs about the linear ordering of individuals along a status continuum. This belief updating is mediated by the medial prefrontal cortex (indicated by a green circle).

observed increased mPFC and amygdala engagement when participants viewed higher-ranked players in an unstable hierarchy (one in which players' positions changed based on their performance). Together, these findings have broadened our understanding of how we represent social hierarchies and their members; however, Kumaran et al. is the first study to formally specify the computational roles of hippocampus and amygdala in the formation of social hierarchy representations. Furthermore, it is the first to directly compare social distance representations from both egocentric and allocentric perspectives.

The neuroimaging results directly comparing representations of one's own versus others' hierarchies demonstrate that one's own social hierarchy is "special," insofar as changes in probabilistic beliefs about self-hierarchy members' status differentially correlated with mPFC activity. This region of mPFC (including pregenual anterior cingulate) is reliably associated with thinking about one's own traits, mental states, and characteristics (Jenkins and Mitchell, 2011). By some accounts, people use self-referential knowledge to make judgments about similar others. For example, one experiment demonstrated repetition suppression in this region of mPFC when participants made judgments about their own preferences following judgments of a similar person's preference (but not when following a dissimilar person's preference; Jenkins et al., 2008). These findings suggest that mPFC may be recruited in service of representing similar others. However, Kumaran et al.'s findings highlight a boundary condition on this proposition: mPFC should only support representations of similar others—including a close friend—when using the self as a template or an anchor is diagnostic (Tamir and Mitchell, 2010). In Kumaran et al.'s experiment, the self is irrelevant to updating beliefs about a hierarchy to which one doesn't belong.

Finally, the results from Kumaran et al.'s categorization task indicate that social hierarchy information is retrieved spontaneously even when the task does not require it. These findings dovetail nicely with research on status in social psychology. Many social comparisons—"Is this person smarter? Better looking?"—happen quickly, consume few cognitive resources, and appear to occur outside of our control (Suls and Wheeler, 2000). This suggests that once hierarchies are established, people will automatically retrieve representations of others' status, with one caveat: not *all* comparisons are diagnostic. Thinking about Warren Buffet's 2016 income presumably has little impact on your assessment of yourself; however, learning that your similarly ranked colleague received a larger raise than you last year might put your teeth on edge. Festinger (1954) suggested that self-evaluations are often derived from social comparisons with people who are self-relevant (i.e., share similar attributes, characteristics, or contexts). Yet other evidence suggests that people generate social comparisons even when they are not diagnostic; they just don't use the comparisons to update their beliefs about themselves (Gilbert et al., 1995).

While social status is a central preoccupation of humans, it has not until recently been a central preoccupation of neuroscientific inquiry. The study reported by Kumaran et al. is a fascinating example of how computational theory, behavior, and brain imaging can together offer insight into how we discover our place in the social world.

REFERENCES

Behrens, T.E.J., Woolrich, M.W., Walton, M.E., and Rushworth, M.F.S. (2007). Nat. Neurosci. *10*, 1214–1221.

De Martino, B., Fleming, S.M., Garrett, N., and Dolan, R.J. (2013). Nat. Neurosci. *16*, 105–110.

Festinger, L. (1954). Hum. Relat. *7*, 117–140.

Fiske, A.P. (1992). Psychol. Rev. *99*, 689–723.

Gilbert, D.T., Giesler, R.B., and Morris, K.A. (1995). J. Pers. Soc. Psychol. *69*, 227–236.

Jenkins, A.C., and Mitchell, J.P. (2011). Soc. Neurosci. *6*, 211–218.

Jenkins, A.C., Macrae, C.N., and Mitchell, J.P. (2008). Proc. Natl. Acad. Sci. USA *105*, 4507–4512.

Kumaran, D., Banino, A., Blundell, C., Hassabis, D., and Dayan, P. (2016). Neuron *92*, this issue, 1135–1147.

Suls, J., and Wheeler, L. (2000). Handbook of Social Comparison: Theory and Research (Springer).

Tamir, D.I., and Mitchell, J.P. (2010). Proc. Natl. Acad. Sci. USA *107*, 10827–10832.

Tavares, R.M., Mendelsohn, A., Grossman, Y., Williams, C.H., Shapiro, M., Trope, Y., and Schiller, D. (2015). Neuron *87*, 231–243.

Thomsen, L., Frankenhuis, W.E., Ingold-Smith, M., and Carey, S. (2011). Science *331*, 477–480.

Zink, C.F., Tong, Y., Chen, Q., Bassett, D.S., Stein, J.L., and Meyer-Lindenberg, A. (2008). Neuron *58*, 273–283.