

Forward Into the Past, Part 2

March 25, 2020



*This column was completed before COVID-19 became a pandemic.
More on that topic next month.*

In [last month's](#) *Observer* column, I suggested that a robust and replicable psychological science might be more achievable if we heed some advice from William James. He advised his 19th-century colleagues to think of psychological categories, such as emotion categories, not as “eternal and sacred psychic entities” (James, 1892, p. 374–375)—i.e., not as psychological types—but as populations of situated, variable instances, similar to Darwin’s understanding of animal species. In this month’s column, let’s follow James’s line of thinking a little further. It leads to one of the most radical but important ideas in psychological science.

After describing emotion categories as populations of variable instances, James went on to write, “[If] we regard them as products of more general causes (as ‘species’ are now regarded as products of heredity and variation), the mere distinguishing and cataloguing becomes of subsidiary importance” (James, 1892, p. 375). James was suggesting that emotional events are created by processes that cross-cut traditional categories of western folk psychology, in the same way that animal species are the products of species-general processes. Some ensembles of genes are species-specific, of course, as agents of heredity. But the *processes* that create species cut *across* species. So, when it comes to psychological science, we

should be focusing on the common processes that, together, create the variable instances of psychological categories. “A science of the relations of mind and brain must show how the elementary ingredients of the former correspond to elementary functions of the latter” (James, 1890/2007, p. 28).

[See all columns by Lisa Feldman Barrett](#)

Over the past century, psychological science started with categories of western folk psychology and tried to map them to measurements of the brain. A more robust scientific approach reverses this ordering, however. It begins with research on the structure and function of the brain, and discovers the means by which the brain produces mental events and actions. This approach points directly to hypotheses about domain-general ingredients of the mind. I'll offer a couple of examples, not as definitive claims, but to illustrate the approach.

The first example comes from a family of interrelated research programs referred to as predictive processing. When considered together, these research programs form a coherent, neurobiologically-inspired research framework united by a common core hypothesis: Your experience of the world and the actions you take derive from an active, constructive process driven by past events. Your brain continuously re-implements (i.e., “remembers”) trajectories of prior events, and in so doing predicts what's going to happen next; those predictions are confirmed or corrected by incoming sense data from the world (e.g., Clark, 2013; Friston, 2010; Hutchinson & Barrett, 2019) and the body (e.g., Hutchinson & Barrett, 2019; Kleckner et al., 2017). Predictive processing hypotheses reverse the causality found in psychology's traditional scientific approach, in which mental events (e.g., thoughts, feelings) and actions (e.g. behaviors), are reactions to sensory inputs from the world (i.e., stimuli); past experiences, if they are relevant, modulate these stimulus-response links.

Predictive processing is not a new idea—it was anticipated by Helmholtz's idea of unconscious inference and has been proposed numerous times in psychological science during the last century (see Box 1, Hutchinson & Barrett, 2019). Nonetheless, this newest generation of predictive processing research does offer something new: the inescapable implication that thoughts, feelings, perceptions, actions, and even hallucinations and delusions are constructed via the same brain-based causes; they are not biological types of psychological phenomena, each implemented in their own specific set of neurons.

[See all APS Presidential Columns since 1999](#)

We still have much to learn about how a brain implements and corrects its predictive dynamics to control action and create mental events, but in broad terms, predictive processing suggests candidates for James's basic ingredients of the human mind: prediction signals (i.e., “memory”) and unexpected sense data from the body and from the world that is encoded to update prediction signals, called prediction error (i.e., “learning”). In this view, “memory” and “learning” are not separate types of mental events, but ongoing processes that are involved, to a greater or lesser extent, in *every* action and mental event. Another ingredient of the mind might be the neural modulation of prediction and prediction error signals (i.e., “executive control”); this ingredient is also thought to be continually present, to some degree, regardless of whether thoughts, feelings, and behavioral responses feel automatic or effortful (e.g., Barrett, Tugade, & Engle, 2004).

And here's another counterintuitive hypothesis: A predictive processing approach further suggests that many psychological phenomena with different names, which we now treat as distinct and separate, may actually be the domain-general mental ingredients. For example, what we refer to as "memory" may be identical to "perceptual inference," "simulation," "ad hoc concept construction," and even "categorization" (e.g., Barrett, 2017). The extravagant assortment of psychological constructs may be ontologically reducible to many fewer mental ingredients.

Analyses of functional magnetic resonance imaging (fMRI) data, particularly during moments when participants are not performing a task (referred to as "resting state" periods), reinforce the hypothesis that "memory," "learning," and "executive control" are possible candidates for James's basic ingredients of the mind. This research has revealed neural "communities," or subnetworks, such as the so-called default mode network, the salience network, and the executive control network that participate in a wide variety of tasks (e.g., Yeo et al., 2015) and dynamic configurations (e.g., Allen et al., 2014). These subnetworks organize along several larger-scale gradients (Margulies et al., 2016; Zhang et al., 2019). One functional gradient describes a spectrum from "re-implementing past experience" (i.e., "memory") at one end to "representing unexpected sense data" (i.e., "learning") at the other. A second gradient runs from representing predictions and prediction errors at one end to modulating those representations (i.e., "executive control") at the other.

Since I'm already way out on a limb here, let's travel a few more inches: your body is also a basic ingredient of your mind. Not in some gauzy metaphorical way, but in a very real, biological way. Neuroscientists have known for some time that a core task of a brain—as a whole system—is to anticipate the body's metabolic needs, attempting to meet those needs before they arise (Sterling & Laughlin, 2015). This process is called allostasis (Sterling, 2012). A brain's estimation of its body's state, in an effort to maintain allostasis, is likely at the core of all mental activity, even in moments that are not bubbling with emotion (Hutchinson & Barrett, 2019; Kleckner et al., 2017).

Since I'm already way out on a limb here, let's travel a few more inches: your body is also a basic ingredient of your mind. Not in some gauzy metaphorical way, but in a very real, biological way.

Many neuroscientists also now agree that sensory systems regulate and are therefore entwined with motor systems, an insight which suggests another mental ingredient: affect. Here's the convoluted logic: If allostasis is a core function of the brain, then so are the sensory consequences of allostasis, called interoception (note that I am not defining interoception as the awareness of sensory signaling from the body, but rather the brain's estimation of the body's metabolic state). Neuroscience research consistently shows that the continuous torrent of interoception is somehow related to a continuous ebb and flow of affective feelings—pleasure and displeasure, comfort and discomfort, arousal and quiescence. The corresponding hypothesis, then, is that affect may be yet another of the mind's ingredients.

Of course, there are many, many challenges to following James's advice. Several have probably already occurred to you, including: Appealing to the basic functions of the brain to discover the basic ingredients of the mind ignores the fact that our notions of how a brain works are inextricably entwined with western concepts of mental life. And the words that name the proposed ingredients of the mind—"memory," "learning," "attention," and "affect"—each refer to a variety of phenomena. For example, "memory" is

used to refer to the implementation of a past experience (to the act of “remembering”), but also a subjective experience of recollection, familiarity, and nostalgia. The search for James’s ingredients of the mind will require that we develop a scientific lingua franca for describing how the brain works separately from what a mind is. We may have to reclaim some words and scientifically redefine them, and then proceed to use them in a consistent way. Otherwise, we are at risk of unwittingly replacing one set of folk psychology categories with another.

If James is right, and “species” of folk psychology categories are actually variable populations emerging from the continual interaction of domain-general processes, then his vision offers us unprecedented scientific opportunities. First, we have the opportunity to assemble a more cumulative science of the mind. The search for domain-general ingredients treats the boundaries between folk categories of perceptions, cognitions, emotions, and actions as *subjectively experienced* (Barrett, 2009; James, 1890/2007, p. 195). We can extend this insight to question the boundaries between the categories of mental disorder, neurodegenerative disorder, and physical disorder. The search for types, a.k.a., a typological mindset, by contrast, shores up those boundaries. Psychological scientists typically search for domain-specific (i.e., type-specific) processes; for example, each psychological type (such as the category “fear”) is presumed to have a strong correspondence to a specific process (e.g., a fear process) and biological mechanism (i.e., a fear circuit). A domain-general approach questions the presumed parallelism between what a mind is, what a mind does, and how a mind is caused (a.k.a., the computational, algorithmic, and implementational levels of analysis, per the trichotomy of neuroscientist David Marr; Marr, 1982). The result, I expect, will be a major course correction in the scientific questions we ask and how we interpret our research findings.

James’s vision also offers us the opportunity to build a generalizable and ultimately universally applicable psychological science. Ingredients of the mind do more than redescribe the brain’s functions in psychological terms—they offer a path to discover how a single, complex, human brain architecture, in continuous conversation with a human body and the world, produces a *variety* of human minds. The minds in other cultures do not necessarily carve their mental lives into perceptions, thoughts, feelings, and actions (see Danziger, 1997, chapter 1). Human brains are responsive to the contexts that humans find themselves in (and help to shape). We need a psychological science that works to describe and explain the diversity of human minds on this planet. A psychological science which does not assume that the kinds of human minds we now observe have always existed. Or will always continue to exist. α

References

- Allen, A. E., et al. (2014). Tracking whole-brain connectivity dynamics in the resting state. *Cerebral Cortex*, 24, 663-676.
- Barrett, L. F. (2009). The future of psychology: Connecting mind to brain. *Perspectives in Psychological Science*, 4, 326-339.
- Barrett, L. F. (2017). [The theory of constructed emotion: An active inference account of interoception and categorization](#). *Social Cognitive and Affective Neuroscience*, 12, 1-23.
- Barrett, L. F., Tugade, M. M., & Engle, R. W. (2004). Individual differences in working memory capacity and dual-process theories of the mind. *Psychological Bulletin*, 130, 553-573.

Clark, A. (2013). Whatever next? Predictive brains, situated agents, and the future of cognitive science. *Behavioral and Brain Sciences*, 36, 281–253.

Danziger, K. (1997). *Naming the mind: How psychology found its language*. London, England: Sage.

Friston, K. (2010). The free-energy principle: A unified brain theory? *Nature Reviews Neuroscience*, 11, 127–38.

Hutchinson, J. B. & Barrett, L. F. (2019). The power of predictions: An emerging paradigm for psychological research. *Current Directions in Psychological Science*, 28, 280–291.

James, W. (1890/2007). *The principles of psychology*. New York: Dover.

James, W. (1892). *Psychology: Briefer course*. New York: Henry Holt.

Kleckner, I. R., et al. (2017). [Evidence for a large-scale brain system supporting allostasis and interoception in humans](#). *Nature Human Behavior*, 1, 0069

Margulies, D. S. et al. (2016). Situating the default-mode network along a principal gradient of macroscale cortical organization. *Proceedings of the National Academy of Sciences*, 113, 12574–12579

Marr, D. (1982). *Vision: A computational investigation into the human representation and processing of visual information*. New York: Freeman.

Sterling, P. (2012). Allostasis: A model of predictive regulation. *Physiology & Behavior*, 106, 5–15.

Sterling, P., & Laughlin, S. (2015). *Principles of neural design*. Cambridge, MA: MIT Press.

Yeo, B. T. et al. (2015). Functional specialization and flexibility in human association cortex. *Cerebral Cortex*, 25, 3654–3672.

Zhang, J., et al (2019). Intrinsic functional connectivity is organized as three interdependent gradients. *Scientific Reports*, 9, 15976