

# THE *Current*

September 7, 2023

[Harrison Tasoff](#)

## **Moral reasoning displays characteristic patterns in the brain, with distinctions between moral categories**

Every day we encounter circumstances we consider wrong: a starving child, a corrupt politician, an unfaithful partner, a fraudulent scientist. These examples highlight several moral issues, including matters of care, fairness and betrayal. But does anything unite them all?

Philosophers, psychologists and neuroscientists have passionately argued whether moral judgments share something distinctive that separates them from non-moral matters. Moral monists claim that morality is unified by a common characteristic and that all moral issues involve concerns about harm. Pluralists, in contrast, argue that moral judgments are more diverse in nature.

Fascinated by this centuries-old debate, a team of researchers set out to probe the nature of morality using one of moral psychology's most prolific theories. The group, led by UC Santa Barbara's [René Weber](#), intensively studied 64 individuals via surveys, interviews and brain imaging on the wrongness of various behaviors.

They discovered that a general network of brain regions was involved in judging moral violations, like cheating on a test, in contrast with mere social norm violations, such as drinking coffee with a spoon. What's more, the network's topography overlapped strikingly with the brain regions involved in theory of mind. However, distinct activity patterns emerged at finer resolution, suggesting that the brain processes different moral issues along different pathways, supporting a pluralist view of moral reasoning. The results, published in [Nature Human Behaviour](#), even reveal differences between how liberals and conservatives evaluate a given moral issue.

"In many ways, I think our findings clarify that monism and pluralism are not necessarily mutually exclusive approaches," said first author [Frederic Hopp](#), who led the study as a doctoral student in UC Santa Barbara's Media Neuroscience Lab. "We show that moral judgments of a wide range of different types of morally relevant behaviors are instantiated in shared brain regions."

That said, a machine-learning algorithm could reliably identify which moral category, or "foundation," a person was judging based on their brain activity. "This is only possible because moral foundations elicit distinct neural activations," Hopp explained.

The group was guided by Moral Foundations Theory (MFT), a framework for explaining the origins and variation in human moral reasoning. "MFT predicts that humans possess a set of innate and universal moral foundations," Weber explained. These are generally organized into six categories:

1. Issues of care and harm,
2. Concerns of fairness and cheating,
3. Liberty versus oppression,
4. Matters of loyalty and betrayal,
5. Adherence to and subversion of authority,
6. And sanctity versus degradation.

The framework arranges these foundations into two broad moral categories: care/harm and fairness/cheating emerge as "individualizing" foundations that primarily serve to protect the rights and freedoms of individuals. Meanwhile loyalty/betrayal, authority/subversion and sanctity/degradation form "binding" foundations, which primarily operate at the group level.

The researchers created a model based on MFT to test whether the framework — and its nested categories — was reflected in neural activity. Sixty-four participants rated short descriptions of behaviors that violated a particular set of moral foundations, as well as behaviors that simply went against conventional social norms, which served as a control. An fMRI machine monitored activity across different regions of their brains as they reasoned through the vignettes.

Certain brain regions distinguished moral from non-moral judgment across the board, such as activity in the medial prefrontal cortex, temporoparietal junction and posterior cingulate, among other regions. Participants also took longer to rate moral transgressions than non-moral ones. The delay suggests that judging moral issues may involve a deeper evaluation of an individuals' actions and how they relate to one's own values, the authors said.

"Although moral judgments are intuitive at first, deeper judgment requires responses to the six 'W questions,'" said senior author Weber, director and lead researcher of [UCSB's Media Neuroscience Lab](#), and a professor in the Departments of Communication and of Psychological and Brain Sciences. "Who does what, when, to whom, with what effect, and why. And this can be complex and takes time." Indeed, moral reasoning recruited regions of the brain also associated with mentalizing and theory of mind.

The researchers also found that transgressions of loyalty, authority and sanctity prompted greater activity in regions of the brain associated with processing other people's actions, as opposed to the self. "It was surprising to us how well the organization into 'individualizing' versus 'binding' moral foundations is reflected on the neurological level in multiple networks," Weber said.

Next, the authors developed a decoding model that accurately predicted which specific moral foundation or social norm individuals were judging from fine-grained activity pattern across their brains. This would not have been possible if all moral categories were unified at the neurological level, they explained.

"This supports MFT's prediction that each moral foundation is not encoded in a single 'moral hotspot,'" the authors write, "but (is instead) instantiated via multiple brain regions distributed across the brain." This finding suggests that the distinct

moral categories proposed by Moral Foundations Theory have an underlying neurologic basis.

In this way, moral reasoning is similar to other mental tasks: it elicits characteristic patterns across the brain, with nuances based on the specifics. For instance, looking at pictures of houses and faces activates a brain region known as the ventral temporal cortex. “However, when looking at the pattern of activation in this region, one can clearly discern whether someone is looking at a house or a face,” Hopp explained. Analogously, moral reasoning activates certain regions of the brain, “yet, the activation patterns in those same regions are highly distinct for different classes of moral behaviors, suggesting that they are not unified.”

Far from merely an esoteric exercise, MFT provides a robust framework for understanding group identity and political polarization. Mounting evidence from survey and behavioral experiments suggests that liberals (progressives) are more sensitive to the categories of care/harm and fairness/cheating, which primarily protect the rights and freedoms of individuals. Conservatives, in contrast, place greater emphasis on the loyalty/betrayal, authority/subversion, and sanctity/degradation categories, which generally operate at the group level.

“Indeed, our results provide evidence at the neurological level that liberals and conservatives have complex differential neural responses when judging moral foundations,” Weber explained. That means individuals at different points along the political spectrum likely emphasize completely different values when evaluating a particular issue.

This paper is part of an avenue of research that the Media Neuroscience Lab embarked on in 2016, aiming to understand how humans make moral judgments, and how the underlying processes vary across more and less realistic scenarios. “the observation that we can reliably decode which moral violation an individual is perceiving also opens exciting avenues for future research: Can we also decode if a moral violation is detected when reading a news story, listening to a radio show, or even when watching a political debate or movie?” Hopp said. “I think these are fascinating questions that will shape the next century of moral neuroscience.”

The study’s co-investigators include renowned neuroscientist and moral philosopher Walter Sinnott-Armstrong from Duke University and [Scott Grafton](#), a professor in UC

Santa Barbara's Department of Psychological and Brain Sciences. [Jacob Fisher](#) and Ori Amir also contributed as co-authors, and were, respectively, a Ph.D. student and a postdoctoral fellow in Weber's Media Neuroscience Lab at the time the work was conducted.

Ultimately, the researchers say, our ability to cooperate in groups is guided by systems of moral and social norms, and the rewards and punishments that result from adhering to or violating them. "For millennia, fables and fairy tales, nursery rhymes, novels, and even 'the daily news' all weave a tapestry of what counts as good and acceptable or as bad and unacceptable," Weber said. "Our results contribute to a better understanding of what moral judgments are, how they are processed, and how they can be predicted across different groups."

Media Contact

**Harrison Tasoff**

Science Writer

(805) 893-7220

[harrisontasoff@ucsb.edu](mailto:harrisontasoff@ucsb.edu)

---

## **About UC Santa Barbara**

The University of California, Santa Barbara is a leading research institution that also provides a comprehensive liberal arts learning experience. Our academic community of faculty, students, and staff is characterized by a culture of interdisciplinary collaboration that is responsive to the needs of our multicultural and global society. All of this takes place within a living and learning environment like no other, as we draw inspiration from the beauty and resources of our extraordinary location at the edge of the Pacific Ocean.