MARCH/APRIL 2020 | MIND.SCIENTIFICAMERICAN.COM

SCIENTIFIC AMERICAN MIND

A CONSCIOUS UNIVERSE?

Some theorize that consciousness is not something unique to humans but is instead a quality inherent to all matter

INCLUDING

PSYCHEDELIC RESEARCH GETS SERIOUS

CLEAR SEX DIFFERENCES IN PERSONALITY

HOW DISINFORMATION HACKS YOUR BRAIN

WITH COVERAGE FROM nature

FROM THE EDITOR

LIZ TORMES

All That We Might Possibly Know

One of my advisers in graduate school used to say that what we humans know about the universe and our existence is a paltry fraction of all that is possible to know. I found this equally tantalizing and frustrating and, like so many other scientists, took comfort in the process of science: a way of thinking that helps you narrow down, through experimentation, observation and critical thinking, what is indeed known. But that nagging truth is still there—that we simply know a lot less than we can ever comprehend as a species. Some areas of research butt up against this reality more than others. In our cover story, journalist Gareth Cook speaks to philosopher Philip Goff on the nature of consciousness. It's a quality that is not measurable by any scientific tool we possess, and so, for now, it lives in the realm of "unknowable." Yet Goff is able to outline an alternative perspective on consciousness that may give us a different vantage point on our own experience (see "Does Consciousness Pervade the Universe?").

Speaking of alternative perspectives, health editor Tanya Lewis details in this issue new research examining the therapeutic potential of psychedelics such as magic mushrooms (see "Giving Psychedelics the Serious Treatment"). And our columnist Scott Barry Kaufman takes an unflinching look at a growing pile of data showing stark personality differences between the sexes. Such differences have long been tiptoed around to avoid controversy (see "Taking Sex Differences in Personality Seriously"). Indeed, sometimes what we discover in science can make us uncomfortable. But overall, isn't it better to know?

Andrea Gawrylewski
Senior Editor, Collections
editors@sciam.com

SCIENTIFIC AMERICAN
MIND

Your Opinion Matters!

Help shape the future of this digital magazine. Let us know what you think of the stories within these pages by emailing us: editors@sciam.com.

GETTY IMAGES

On the Cover

What if consciousness is not something special that the brain does but is instead a quality inherent to all matter?

WHAT'S INSIDE

March-April 2020
Volume 31 • Number 2

GETTY IMAGES

MOHA ELJAW GETTY IMAGES

NEWS

4.
Americans Are Fast to Judge Social Class
Judgments about the way people talk happen quickly and affect hiring decisions

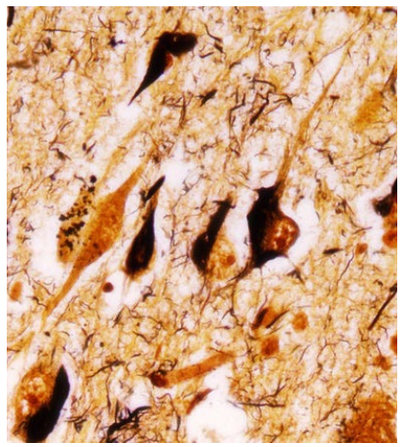
6.
The Brain Senses Touch beyond the Body
You detect a tool's contact with an object as if you placed your own finger on it

7.
Possible Missing Link in Alzheimer's Pathology Identified
It may open the door to new treatments and explain why previous ones failed

9.
Emotional Words Such as "Love" Mean Different Things in Different Languages
An analysis of more than 2,000 languages reveals differences in the way feelings are conceptualized among cultures

11.
Scientists Spot Addiction-Associated Circuit in Rats
Rats show changes in compulsive behavior when a brain connection is turned on or inhibited

MARIA DUBOVA GETTY IMAGES



GETTY IMAGES

FEATURES

13.
Does Consciousness Pervade the Universe?
Philosopher Philip Goff answers questions about "panpsychism"

16.
Johns Hopkins Scientists Give Psychedelics the Serious Treatment
The first research center of its kind in the country is bringing renewed rigor to the investigation of the drugs' therapeutic uses

19.
Taking Sex Differences in Personality Seriously
New approaches are shedding light on the magnitude of sex differences in personality, and the results are so strong and pervasive that they can no longer be ignored

26.
What's Next for Psychology's Embattled Field of Social Priming
A promising field of research on social behavior struggled after investigators couldn't repeat key findings. Now researchers are trying to establish what's worth saving

OPINION

31.
How Disinformation Hacks Your Brain
The digital age has heightened our vulnerability to falsehood, but recognizing such weaknesses can help guard against them

33.
The Language You Speak Influences Where Your Attention Goes
It's all because of the similarities between words

35.
How a Flawed Experiment "Proved" That Free Will Doesn't Exist
It did no such thing—but the result has become conventional wisdom nonetheless

NEWS

CECILE ARCURS GETTY IMAGES

Americans Are Fast to Judge Social Class

Judgments about the way people talk happen quickly and affect hiring decisions

The plot of the famous musical *My Fair Lady* is based on the idea that the way we speak determines our position in society. The main character, Eliza Doolittle, becomes the unwitting target of a bet between two phonetics scholars, one of whom (Henry Higgins) brags that he can convince strangers that Doolittle is a duchess by training her to speak like one. In reality, she is the poor daughter of a dustman who speaks with a thick Cockney accent. By the end of the musical, Doolittle is able to pronounce all of her words like a member of the British elite, fooling everyone at an embassy ball about her true origins.

Based on a new set of scientific studies, it seems that Higgins may have been right: people can determine our social class by the way we talk. Michael Kraus and his colleagues at Yale University recently published a paper in the *Proceedings of the National Academy of Sciences USA* entitled “Evidence for the

Reproduction of Social Class in Brief Speech." The paper lays out evidence from five studies demonstrating that people can accurately judge someone's social standing from that individual's speech and that people use these judgments to discriminate against lower-class job candidates.

It's hard to imagine a version of *My Fair Lady* set in the U.S. because, unlike the British, Americans seem either unwilling or unable to honestly acknowledge their own social class. A 2015 poll by the Pew Research Center found that the majority of Americans consider themselves broadly "middle class," whether they are making less than $30,000 or more than $100,000 per year. But as the new research demonstrates, Americans find it easy to make distinctions about other people's social class just by listening to them speak.

In one study, Kraus and his colleagues asked 229 people to listen to 27 different speakers who varied in terms of their age, race, gender and social class. The study participants heard each speaker say a total of seven different words. Based on just this short audio, participants were able to correctly identify which speakers were college-educated 55 percent of the time—more than what would be expected by chance. A major limitation of this study, however, was that it used college education as a proxy for social class. In addition, the researchers wanted to examine the hypothesis that people infer social class from speaking style rather than the content of what is said.

Therefore, in another study, they ran an experiment where 302 participants were asked to either listen to or read transcripts from 90 seconds of recorded speech in which the speakers talked about themselves without explicitly mentioning anything about their social class (for example, their job title). Participants were asked to judge what they thought the social classes of the speakers were by using a 10-rung ascending ladder of increasing income, education and occupation status. They found that participants who heard the audio recordings were more accurate in judging where the speakers fell in terms of their social status. This finding suggests that we infer people's social class largely from how they talk rather than what they say.

To demonstrate whether these inferences have real-world consequences, Kraus and his colleagues ran another experiment in the form of a simulated hiring scenario. They recruited 20 prospective job candidates from a pool of 110 applicants to practice interviewing for a laboratory manager position requiring a broad range of technical and interpersonal skills. The 20 candidates were chosen because they represented the widest disparity between high and low social class from the entire applicant pool. Each candidate was video recorded while answering the question "How would you describe yourself?" The researchers recruited 274 participants, all of whom had past hiring experience, to either listen to the audio from these videos or read a transcript of the content.

The findings showed that participants were able to accurately judge the social class of the candidates and that this effect was stronger for participants who had heard the audio recordings. In addition, participants judged the higher-class candidates as more competent, a better fit for the job and more likely to be hired. They also awarded them a higher starting salary and a larger sign-on bonus.

Taken together, this research suggests that despite our discomfort about the topic, Americans are able to easily detect one another's social class from small snippets of speech. Moreover, we use this information to discriminate against people who seem to be of a lower social class. Most of us are aware that employment laws protect us from being unfairly discriminated against for characteristics beyond our control, such as gender or race. This research identifies social class as another potential way that employers may discriminate against candidates, perhaps without even realizing it.

Certainly there is a lot more research that needs to be done before we can draw firm conclusions about how social class impacts discrimination. For example, it would be useful to understand how stable people's speech patterns are over time and after exposure to different situations. In addition, researchers could test whether making hiring managers more aware of social-class bias changes their judgments about candidates. The hope is that this paper will spur more scientists to pay attention to the ways in which speech plays a fundamental role in creating and maintaining social inequality.

—Daisy Grewal

The Brain Senses Touch beyond the Body

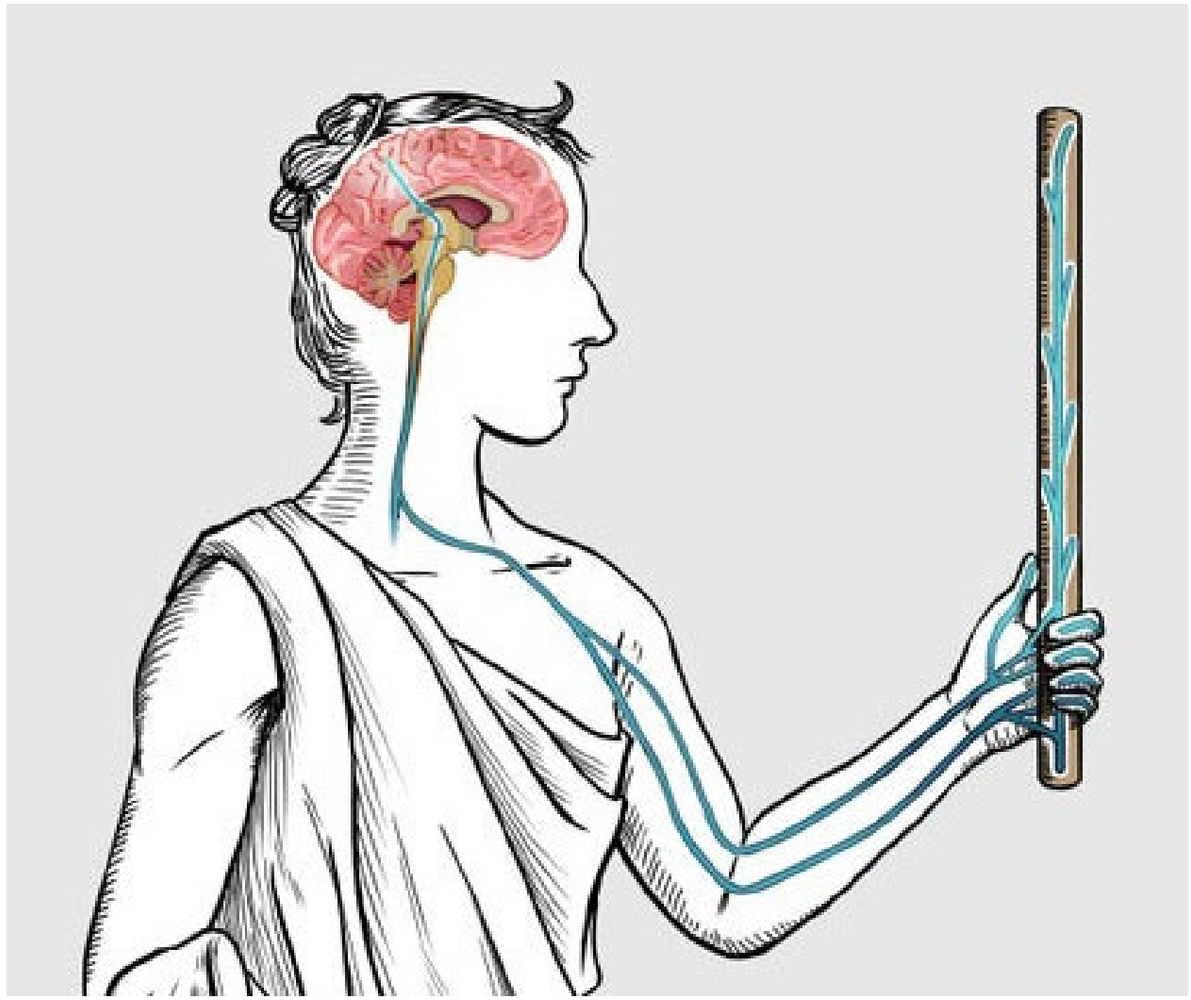
You detect a tool's contact with an object as if you placed your own finger on it

Luke Miller, a cognitive neuroscientist, was toying with a curtain rod in his apartment when he was struck by a strange realization. When he hit an object with the rod, even without looking, he could tell where it was making contact like it was a sensory extension of his body. "That's kind of weird," Miller recalls thinking to himself. "So I went [to the lab], and we played around with it in the lab."

Sensing touch through tools is not a new concept, although it has not been extensively investigated. In the 17th century, philosopher René Descartes discussed the ability of blind people to sense their surroundings through their walking cane. While scientists have researched tool use extensively, they typically focused on how people move the tools. "They, for the most part, neglected the sensory aspect of tool use," Miller says.

In a 2018 *Nature* study, Miller and his colleagues at Claude Bernard Lyon 1 University in France reported that humans are actually quite good at pinpointing where an object comes into contact with a handheld tool using touch alone, as if the object were touching their own skin. A tool is not innervated like our skin, so how does our brain know when and where it is touched? Results in a follow-up study, published last December in *Current Biology,* reveal that the brain regions involved with sensing touch on the body similarly process it on the tool. "The tool is being treated like a sensory extension of your body," Miller says.

In the initial experiment, the researchers asked 16 right-handed subjects to determine where they felt touches on a one-meter-long wooden rod. In a total of 400 trials, each subject compared the locations of two touches made on the rod: If they were felt in different locations, participants did not respond. If they were in the same location, the people in the study tapped a foot pedal to indicate whether the touches were close or far from their hand. Even without any experience with the rod or feedback on their performance, the participants were, on average, 96 percent accurate.



Somatosensory system, running from brain to hand, extends imaginatively outward and into the stick the woman holds.

LOTTIE CLARK

During the experiment, researchers recorded subjects' cortical brain activity using scalp electrodes and found that the cortex rapidly processed where the tool was touched. In trials in which the rod was touched in the same location twice in a row, there was a marked suppression of neural responses in brain areas previously shown to identify touch on the body, including the primary somatosensory (touch) cortex and the posterior parietal cortex.

There is evidence that when the

sensory brain regions are presented with the same stimulus repeatedly, the responses of the underlying neural population gets suppressed. This repetition suppression can be measured and used as a "time stamp" to signify when a stimulus is extracted in the brain.

When the team tested some of the same subjects with touches on their arm instead of the rod, it observed similar repetition suppression in the same brain regions on similar time scales. The somatosensory cortex was suppressed in 52 milliseconds (about one 20th of a second) after contact on both the rod and the arm. At 80 milliseconds, that activity suppression spread throughout the posterior parietal cortex. These results indicate the neural mechanisms for detecting touch location on tools "are remarkably similar to what happens to localize touch on your own body," says Alessandro Farnè, a neuroscientist at the Lyon Neuroscience Research Center in France and senior author of both studies.

Interestingly, after each contact, the rod vibrates for about 100 milliseconds, Miller says. "So by the time the rod is done vibrating in the hand, you've already extracted the location dozens of milliseconds before that," he adds. The vibrations on the rod are detected by touch sensors embedded in our skin called Pacinian receptors, which then relay neural signals up to the somatosensory cortex. Computer simulations of Pacinian activity in the hand showed that information about rod contact location could be extracted efficiently within 20 milliseconds.

The vibrations on the rod may provide the key information needed for touch localization. Repeating the same rod experiment, the researchers tested a patient who lost proprioception in her right arm, meaning she could not sense the limb's location in space. She could still sense superficial touch, however, and she was able to localize where the rod was touched when held in both hands and had similar brain activity as the healthy patients during the task. That finding "suggests quite convincingly that vibration conveyed through the touch, which is spared in the patient, is sufficient for the brain to locate touches on the rod," Farnè says.

Taken together, these results indicate that people could locate touches on a tool quickly and efficiently using the same neural processes for detecting touch on the body. While Farnè emphasizes that no one in the studies thought the tool had "become part of their own body," he says the work indicates the subjects experienced sensory embodiment, "in which the brain repurposes strategies for dealing with objects by reusing what it knows about the body."

"This is really beautiful, comprehensive and thoughtful work," says Scott Frey, a cognitive neuroscientist researching neuroprosthetics at the University of Missouri. Frey, who was not involved with the studies, believes that the results could help inform the design of better prostheses because it suggests that "insensate objects can become, potentially, ways of detecting information from the world and relaying it toward the somatosensory systems," he says. "And that's not something that I think people in the world of prosthetics design really thought about. But maybe this suggests that they should. And that's kind of a neat, novel idea that could come out of it."

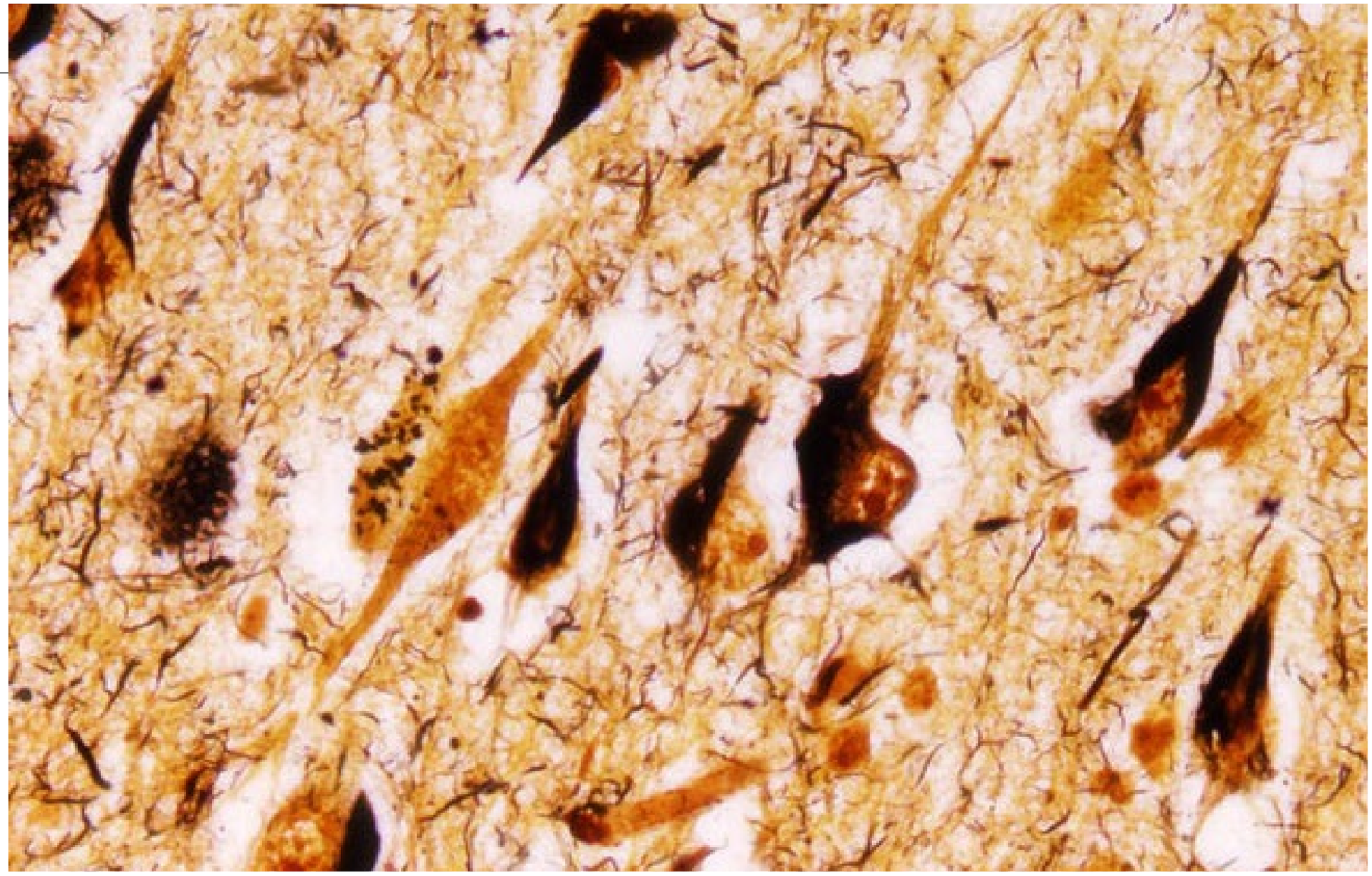
—Richard Sima

Possible Missing Link in Alzheimer's Pathology Identified

It may open the door to new treatments and explain why previous ones failed

Alzheimer's disease has long been characterized by the buildup of two distinct proteins in the brain: first beta-amyloid, which accumulates in clumps, or plaques, and then tau, which forms toxic tangles that lead to cell death. But how beta-amyloid leads to the devastation of tau has never been precisely clear. Now a new study at the University of Alabama at Birmingham appears to describe that missing mechanism.

The study details a cascade of events. Buildup of beta-amyloid activates a receptor that responds to a brain chemical called norepinephrine, which is commonly known for mobilizing the brain and body for action. Activation of this receptor by both beta-amyloid and norepinephrine boosts the activity of an enzyme that activates tau and increases the vulnerability of brain cells to it,



according to the study, published in *Science Translational Medicine*.

Essentially, beta-amyloid hijacks the norepinephrine pathway to trigger a toxic buildup of tau, says Qin Wang, the study's senior author and a professor of neuropharmacology in the department of cell, developmental and integrative biology at the University of Alabama at Birmingham. "We really show that this norepinephrine is a missing piece of this whole Alzheimer's disease puzzle," she says.

This cascade explains why so many previous Alzheimer's treatments have failed, Wang says. Most of the drugs developed in recent decades have targeted the elimination of beta-amyloid. But the new research suggests that norepinephrine amplifies the damage wrought by that protein.

Beta-amyloid itself can kill neurons but only in very high doses, Wang says. Add norepinephrine, and it takes only 1 to 2 percent as much beta-amyloid to eliminate brain cells in a lab dish. So with treatments that targeted beta-amyloid but left the norepinephrine pathway intact, there was enough beta-amyloid remaining to do significant damage, she says. But if the norepinephrine pathway really is crucial to the development of Alzheimer's, it suggests new ways to treat the disease, which currently afflicts 5.8 million Americans.

A drug that was developed to treat depression but was too ineffective to win approval seems to act on this same pathway, Wang says. The drug, idazoxan, which has also been studied in schizophrenia, has already passed through initial clinical testing and been shown to be safe, she adds.

Wang is now looking to promote larger clinical trials of idazoxan to see if it can be used to effectively treat early-stage Alzheimer's. She hopes that eventually, a drug can be developed that will act on this Alzheimer's-related pathway in a more targeted way to minimize side effects and maximize effectiveness.

Stephen Salloway, a professor of psychiatry and neurology at the Warren Alpert Medical School at Brown University, who was not involved in the new research, says he doesn't think Alzheimer's will yield so easily to a new drug targeting the norepinephrine pathway. "I doubt there's something simple that's going to come out of this," says Salloway, who is also director of neurology and the Memory and Aging Program at

GETTY IMAGES

Butler Hospital in Providence, R.I. "I'd be shocked if it works."

Such a drug might, however, be part of a "therapeutic package" of treatments that could eventually make headway against Alzheimer's, he says. "The goal is to get a biological foothold and then build on it," he adds. "The more targets we have, the bigger the impact."

Eric Reiman, CEO of Banner Alzheimer's Institute, an Arizona-based research and advocacy group, agrees that the study suggests new possibilities for treatment. "It provides a mechanism that could be targeted by investigational and, potentially, repurposed drugs," he says. "And it offers hypotheses that can now be tested and extended by the field." Salloway, Reiman and other experts emphasize that the findings are preliminary and need to be confirmed by future research.

Wang has long studied norepinephrine because of its role in thinking and complex behaviors. She stumbled across the connection to Alzheimer's as part of that research, she says.

In two strains of mice and in human tissue in their new study, she and her colleagues showed that small pieces of beta-amyloid bind to a receptor for norepinephrine, activating the GSK3-beta enzyme and triggering the tau to become toxic. They confirmed this relationship by blocking the receptor with idazoxan in two strains of middle-aged mice for eight weeks. Doing so deactivated the enzyme and prevented the tau from becoming toxic.

For years, researchers had wondered how beta-amyloid and tau were connected, says Rudolph Tanzi, an expert on the molecular genetics of Alzheimer's at Massachusetts General Hospital, who was not involved in the new research. Scientists essentially assumed that beta-amyloid caused tau tangles through a complicated chain of events, he says.

Then, in a 2014 paper in *Nature,* Tanzi and his colleagues used human brain cells grown in a dish to reveal a problem with the theory: mice—the main source of research information on Alzheimer's—do not have the right form of tau to cause tangles in people. Instead the researchers showed that in the human cells, beta-amyloid led directly to tau tangles. Tanzi and his colleagues blocked a variety of different enzymes called kinases to try to stop the process. They found two, both of which blocked the GSK3-beta enzyme—the same one that Wang and her colleagues identified in their *Science Translational Medicine* paper.

The new study, Tanzi says, takes his own work a step further by showing how beta-amyloid triggers activation of the toxic tau. "It's an important paper," he adds. "If it's replicated, it provides a good drug target."

Tanzi believes that inflammation is a key player in Alzheimer's, triggering the cascade that leads to disease. He has previously described beta-amyloid as the match and tau tangles as the brushfires burning in the brains of people with the disease. "GSK3-beta, I guess you could say, is the kindling for the brushfire. And this explains how the match gains access to the kindling," Tanzi says. Once the neuroinflammation starts, brain cells die at a far faster rate, he adds.

Tanzi says he has unpublished data on dozens of drugs that stop beta-amyloid from triggering tau tangles, many of which support what Wang and her colleagues found in their new paper. "I believe their data are going to hold up," he says. "And it's exciting."

—Karen Weintraub

Emotional Words Such as "Love" Mean Different Things in Different Languages

An analysis of more than 2,000 languages reveals differences in the way feelings are conceptualized among cultures

Humans boast a rich trove of words to express the way we feel. Some are not easily translatable between languages: Germans use "*Weltschmerz*" to refer to a feeling of melancholy caused by the state of the world. And the indigenous Baining people of Papua New Guinea say "*awumbuk*" to describe a social hangover that leaves people unmotivated and listless for days after the departure of overnight guests. Other terms seem rather common—"fear," for example, translates to "*takot*" in Tagalog and "*ótti*" in Icelandic. These similarities and differences raise a question: Does the way we experience emotions cross cultural boundaries?

Scientists have long questioned whether human emotions share universal roots or vary across cultures. Early evidence suggested that, in the same way that primary colors give rise to all of the other hues, there was a core set of primary emotions from which all other feelings arose. In the 1970s, for instance, researchers reported that people in an isolated cultural group in Papua New Guinea were able to correctly identify emotional expressions in photographed Western faces at rates higher than chance. "This was largely taken as evidence that people around the world could understand emotions in the same way," says Kristen Lindquist, an associate professor of psychology and neuroscience at the University of North Carolina at Chapel Hill.

But more recent studies have challenged this idea. Work from a variety of fields—psychology, neuroscience and anthropology—has provided evidence that the way people express and experience emotions may be greatly influenced by our cultural upbringing. Many of these studies have limitations, however. Most have either looked only at comparisons between two cultures or focused on big, industrialized countries, says Joshua Jackson, a doctoral student in psychology at U.N.C. Chapel Hill. "We haven't really had the power to test [the universality of emotion] on an appropriate scale."

To explore the question of common emotions, Jackson, Lindquist and their colleagues teamed up with researchers at the Max Planck Institute for the Science of Human History in Jena, Germany, in one of the largest studies of cross-cultural emotional expression to date. Their work, which was published last December in *Science*, drew on vocabulary from 2,474 languages. It revealed a great deal of variability in the way emotions are verbally expressed—as well as some underlying commonalities. "Psychologists have been debating whether emotions are universal or variable across cultures for a long time," Jackson says. "I think what this paper shows is that both sides have some merit."

To examine variability in emotional expression, the researchers used computational tools to create a massive database of colexifications, instances where a single word has multiple meanings. Examples include "*ruka*," which means both hand and

MARIA DUBOVA *GETTY IMAGES*

arm in Russian, and "funny," which means both odd and humorous in English. Previous investigations of nonemotional words have demonstrated that colexified ones tend to have common properties—words that describe "sea" and "water" are more likely to be paired than those for "sun" and "water"—suggesting that speakers of a language perceive similarities in them.

The team then used its database to generate networks of colexified words among 20 language families (groups of languages that share ancestral roots) to compare emotion-associated vocabulary worldwide. Doing so revealed significant differences in how emotions were conceptualized across cultures—three times more variation than in terms used to describe color. For example, in some languages, the words for "surprise" tended be

grouped with those for fear, while in others, the same concept was paired with more pleasant states, such as happiness. Through further analysis, the researchers also found that this diversity was partially dependent on the geographical proximity of language families—the closer they were, the more commonalities they were likely to share. "That suggests the extent to which cultures were likely to have historical contact, either via trade or migration or conquest allowed these cultures to interact and perhaps transmit and borrow emotion concepts from one another," Lindquist says.

On the other hand, the researchers also found some underlying similarities. Language families tended to differentiate emotions based on their valence (how pleasant or unpleasant they were) and activation (the level of excitement they elicited). For instance, words that expressed joy were unlikely to be grouped together with those for regret. There were exceptions, however: some Austronesian languages paired the concept of love, a typically positive emotion, with pity, a typically negative one.

"This is an important study," says William Croft, a professor of linguistics at the University of New Mexico, who was not involved in the work. "It's probably the first time an analysis of the meanings of words has been done at this scale." One of the novel things about this project is that the findings show both universal and culture-specific patterns, Croft adds. He points out, however, that because some of these families cover a large number of languages across a wide geographical area, it will be important to further examine the underlying cultural factors.

Another limitation of the study lies in the imperfect nature of translations, says Asifa Majid, a professor of psychology at the University of York in England, who penned an accompanying commentary. This is especially the case when it comes to words for emotion, which can be difficult to express in words—linguists may only obtain approximate translations of such terms while documenting word lists out in the field. Nevertheless, these findings raise a fascinating question about cross-cultural variation in human emotion, Majid adds. "Where we find variation, is it only in language, or is it reflecting something deeper about how people experience emotions, too?"

—Diana Kwon

GETTY IMAGES

Scientists Spot Addiction-Associated Circuit in Rats

Rats show changes in compulsive behavior when a brain connection is turned on or inhibited

For many people battling addictions, seeing drug paraphernalia—or even places associated with past use—can ignite cravings that make relapse more likely. Associating environmental cues with pleasurable experiences is a basic form of learning, but some researchers think such associations can "hijack" behavior, contributing to

problems such as addiction and eating disorders.

Researchers led by neuroscientist Shelly Flagel of the University of Michigan have found a brain circuit that may control this hijacking; rats that exhibit a type of compulsive behavior show different brain connectivity and activity than those that do not, and manipulation of the circuit altered their behavior. These findings may help researchers understand why some individuals are more susceptible to impulse-control disorders. "This is technically a really excellent study," says neuroscientist Jeff Dalley of the University of Cambridge, who was not involved in the work.

In the study, published last September in *eLife,* researchers showed rats an inert lever shortly before delivering a tasty treat via a chute, then sorted them into groups based on their responses. All rats learned to associate the lever with the treat, but some—dubbed "goal trackers"—began to approach the food chute directly after seeing the lever, whereas inherent "sign trackers" kept compulsively returning to the lever itself.

The team suspected that two brain regions were involved: the paraventricular nucleus of the thalamus (PVT), which drives behavior, and the prelimbic cortex, which is involved in reward learning. The researchers used a technique called chemogenetics to alter neurons in the circuit connecting these regions, which let them turn on or inhibit signals from the prelimbic cortex using drugs. Activating the circuit reduced sign trackers' tendency to approach the lever but did not affect goal trackers. Deactivating it drew goal trackers to the lever (sign-tracking behavior), without affecting preexisting sign trackers. The team also found increased dopamine, a chemical messenger involved in reward processing, in the newly sign-tracking brains.

The prelimbic cortex appears to exert top-down control, whereas the PVT processes the motivational signal triggered by the cue. "Individuals seem to be wired differently regarding this balance between top-down cortical control versus bottom-up subcortical processes that are more emotional," Flagel says. Those "who are highly reactive to cues in the environment may suffer from deficits in top-down control." She suggests that cognitive-training therapies might combat such deficits in humans.

The circuit itself could also represent a new treatment target, but the exact human anatomy is unclear, Dalley notes—and addiction is more complex than a single mechanism.

Next, the researchers will try to examine these traits in people. "Once we've established the sign- and goal-tracker paradigm in humans, we can test whether these traits are predictive of psychopathology," Flagel says. "We hope this will help identify individuals who are more susceptible to certain mental illnesses or facets such as relapse."

—*Simon Makin*

Does Consciousness Pervade the Universe?

Philosopher Philip Goff answers questions about "panpsychism"

By Gareth Cook

GETTY IMAGES

Gareth Cook is a Pulitzer Prize–winning journalist who edits *Scientific American*'s Mind Matters online news column.

One of science's most challenging problems is a question that can be stated easily: Where does consciousness come from? In his new book *Galileo's Error: Foundations for a New Science of Consciousness,* philosopher Philip Goff considers a radical perspective: What if consciousness is not something special that the brain does but is instead a quality inherent to all matter? It is a theory known as "panpsychism," and Goff guides readers through the history of the idea, answers common objections (such as "That's just crazy!") and explains why he believes panpsychism represents the best path forward. He answered questions from Mind Matters editor Gareth Cook.

An edited transcript of the interview follows.

Can you explain, in simple terms, what you mean by panpsychism?

In our standard view of things, consciousness exists only in the brains of highly evolved organisms, and hence consciousness exists only in a tiny part of the universe and only in very recent history. According to panpsychism, in contrast, consciousness pervades the universe and is a fundamental feature of it. This doesn't mean that literally everything is conscious. The basic commitment is that the fundamental constituents of reality—perhaps electrons and quarks—have incredibly simple forms of experience. And the very complex experience of the human or animal brain is somehow derived from the experience of the brain's most basic parts.

It might be important to clarify what I mean by "consciousness," as that word is actually quite ambiguous. Some people use it to mean something quite sophisticated, such as self-awareness or the capacity to reflect on one's own existence. This is something we might be reluctant to ascribe to many nonhuman animals, never mind fundamental particles. But when I use the word "consciousness," I simply mean *experience:* pleasure, pain, visual or auditory experience, et cetera.

Human beings have a very rich and complex experience; horses less so; mice less so again. As we move to simpler and simpler forms of life, we find simpler and simpler forms of experience. Perhaps, at some point, the light switches off, and consciousness disappears. But it's at least coherent to suppose that this continuum of consciousness fading while never quite turning off carries on into inorganic matter, with fundamental particles having almost unimaginably simple forms of experience to reflect their incredibly simple nature. That's what panpsychists believe.

You write that you come to this idea as a way of solving a problem in the way consciousness is studied. What, in your mind, is the problem?

Despite great progress in our scientific understanding of the brain, we still don't have even the beginnings of an explanation of how complex electrochemical signaling is somehow able to give rise to the inner subjective world of colors, sounds, smells and tastes that each of us knows in our own case. There is a deep mystery in understanding how what we know about ourselves from the inside fits together with what science tells us about matter from the outside.

While the problem is broadly acknowledged, many people think we just need to plug away at our standard methods of investigating the brain, and we'll eventually crack it. But in my new book, I argue that the problem of consciousness results from the way we designed science at the start of the scientific revolution.

A key moment in the scientific revolution was Galileo's declaration that *mathematics* was to be the language of the new science, that the new science was to have a purely *quantitative* vocabulary. But Galileo realized that you can't capture consciousness in these terms, as consciousness is an essentially *quality-involving* phenomenon. Think about the redness of a red experiences or the smell of flowers or the taste of mint. You can't capture these kinds of *qualities* in the purely quantitative vocabulary of physical science. So Galileo decided that we have to put consciousness outside the domain of science; after

we had done that, everything else could be captured in mathematics.

This is really important because although the problem of consciousness is taken seriously, most people assume our conventional scientific approach is capable of solving it. And they think this because they look at the great success of physical science in explaining more and more of our universe and conclude that this ought to give us confidence that physical science alone will one day explain consciousness. But I believe that this reaction is rooted in a misunderstanding of the history of science. Yes, physical science has been incredibly successful. But it's been successful precisely because it was designed to exclude consciousness. If Galileo were to time travel to the present day and hear about this problem of explaining consciousness in the terms of physical science, he'd say, "Of course, you can't do that. I designed physical science to deal with *quantities,* not *qualities*."

How does panpsychism allow you to approach the problem differently?

The starting point of the panpsychist is that physical science doesn't actually tell us what matter is. That sounds like a bizarre claim at first; you read a physics textbook, you seem to learn all kinds of incredible things about the nature of space, time and matter. But what philosophers of science have realized is that physical science, for all its richness, is confined to telling us about the *behavior* of matter, what it *does*. Physics tells us, for example, that matter has mass and charge. These properties are completely defined in terms of behavior, things like attraction, repulsion, resistance to acceleration. Physics tells us absolutely nothing about what philosophers like to call *the intrinsic nature* of matter: what matter is, in and of itself.

So it turns out that there is a huge hole in our scientific story. The proposal of the panpsychist is to put consciousness in that hole. Consciousness, for the panpsychist, is the intrinsic nature of matter. There's just matter, on this view, nothing supernatural or spiritual. But matter can be described from two perspectives. Physical science describes matter "from the outside," in terms of its behavior. But matter "from the inside"—that is, in terms of its intrinsic nature—is constituted of forms of consciousness.

What this offers us is a beautifully simple, elegant way of integrating consciousness into our scientific worldview, of marrying what we know about ourselves from the inside and what science tells us about matter from the outside.

What are the objections to this idea that you hear most frequently? And how do you respond?

Of course, the most common one is "That's just crazy!" But many of our best scientific theories are wildly counter to common sense, too—for example, Albert Einstein's theory that time slows down when you travel very fast or Charles Darwin's theory that our ancestors were apes. At the end of the day, you should judge a view not by its cultural associations but by its explanatory power. Panpsychism gives us a way of resolving the mystery of consciousness, a way that avoids the deep difficulties that plague more conventional options.

Do you foresee a scenario in which panpsychism can be tested?

There is a profound difficulty at the heart of the science of consciousness: consciousness is unobservable. You can't look inside an electron to see whether or not it is conscious. But nor can you look inside someone's head and see their feelings and experiences. We know that consciousness exists not from observation and experiment but by being conscious. The only way we can find out about the consciousness of others is by asking them: I can't directly perceive your experience, but I can ask you what you're feeling. And if I'm a neuroscientist, I can do this while I'm scanning your brain to see which bits light up as you tell me what you're feeling and experiencing. In this way, scientists are able to correlate certain kinds of brain activity with certain kinds of experience. We now know which kinds of brain activity are associated with feelings of hunger, with visual experiences, with pleasure, pain, anxiety, et cetera.

This is really important information, but it's not itself a theory of consciousness. That's because what we ultimately want from a science of consciousness is an *explanation* of those correlations. Why is it that, say, a certain kind of activity in the hypothalamus is associated with the feeling of hunger? Why should that be so? As soon as you start to answer this question, you move beyond what can be, strictly speaking, tested, simply because consciousness is unobservable. We have to turn to philosophy.

The moral of the story is that we need both the science and the philosophy to get a theory of consciousness. The science gives us correlations between brain activity and experience. We then have to work out the best philosophical theory that explains those correlations. In my view, the only theory that holds up to scrutiny is panpsychism.

How did you become interested in this topic?

When I studied philosophy, we were taught that there were only two approaches to consciousness: either you think consciousness can be explained in conventional scientific terms, or you think consciousness is something magical and mysterious that science will never understand. I came to think that both these views were pretty hopeless. I think we can have hope that we will one day have a science of consciousness, but we need to rethink what science is. Panpsychism offers us a way of doing this.

Giving Psychedelics the Serious Treatment

The first research center of its kind in the country is bringing renewed rigor to the investigation of the drugs' therapeutic uses

By Tanya Lewis

Researchers at Johns Hopkins University are testing whether the potent psychedelic in psilocybin mushrooms can treat everything from smoking addiction to anorexia.

MOHA EL-JAW GETTY IMAGES

Tanya Lewis is an associate editor at *Scientific American* who covers health and medicine.

PSYCHEDELIC DRUGS—once promising research subjects that were decades ago relegated to illicit experimentation in dorm rooms—have been steadily making their way back into the lab for a revamped 21st-century-style look. Scientists are rediscovering what many see as the substances' astonishing therapeutic potential for a vast range of issues, from depression to drug addiction and acceptance of mortality. A frenzy of interest has captivated a new generation of researchers, aficionados and investors, triggering some understandable wariness over promises that may sound a little too good to be true. But late last year the highly respected institution Johns Hopkins University—the U.S.'s oldest research university—launched a dedicated center for psychedelic studies, the first of its kind in the country and perhaps the world's largest. With work now underway, the center is aiming to enforce the strictest standards of scientific rigor on a field that many feel has veered uncomfortably close to mysticism and that has relied heavily on subjective reports. Early results have been promising and seem poised to keep the research on a roll.

Psilocybin (a psychoactive compound found in certain mushrooms) and LSD were widely studied in the 1950s and 1960s as treatments for alcoholism and other maladies. They later gained a reputation in the media and the public eye as dangerous and became strongly associated with the counterculture. Starting in 1966, several states banned their use. In 1968 LSD was outlawed nationwide, and in 1970 Congress passed the Controlled Substances Act, classifying that drug and psilocybin, along with several others, as having a high potential for abuse and no accepted medical use. But in recent years a rapidly growing number of studies reporting encouraging results in treating depression, addiction and post-traumatic stress disorder (PTSD) have brought them back out of the shadows, spurred on by positive media coverage.

In a major boost to the reviving field, Johns Hopkins's Center for Psychedelic and Consciousness Research is exploring the use of psychedelics—primarily psilocybin—for problems ranging from smoking addiction to anorexia and to Alzheimer's disease. "One of the remarkably interesting features of working with psychedelics is they're likely to have transdiagnostic applicability," says Roland Griffiths, who heads the new facility and has led some of the most promising studies evaluating psilocybin for treating depression and alcoholism. The myriad applications suggested for these drugs may be a big part of what makes them sound, to many, like snake oil—but "the data [are] very compelling," Griffiths says. And psychedelics may not only hold hope for treating mental disorders. As Griffiths puts it, they provide an opportunity to "peer into the basic neuroscience of how these drugs affect brain activity and worldview in a way that is ultimately very healthy."

As author Michael Pollan chronicles in his 2018 best seller *How to Change Your Mind*, researchers were examining the therapeutic effects of psychedelics in the 1950s—a decade before then Harvard University psychologist Timothy Leary and his colleague Richard Alpert started their notorious study in which they gave psilocybin to students (ultimately leading to Leary's and Alpert's dismissal from the university). In the 1950s–1970s, studies conducted with LSD—which acts on the same brain receptors as psilocybin—reported strong results in treating substance use disorders, including alcohol and heroin addiction. But when LSD became illegal in 1968, funding for this work gradually dried up. Most psychedelics research stopped or went underground.

PSYCHEDELICS' NEW WAVE

Griffiths and some of his colleagues helped to revive the field around 2000, when they obtained government approval to give high doses of psilocybin to healthy volunteers. The researchers published a foundational study in 2006 showing that a single dose was safe and could cause sustained positive effects and even produce "mystical experiences." A decade later they published a randomized double-blind study demonstrating psilocybin significantly decreased depression and anxiety in patients who had life-threatening cancer. Each participant underwent two sessions (a high-dose one and a low-dose one) five weeks apart. Six months afterward, about 80 percent of the patients were still less clinically depressed and anxious than before undergoing the treat-

ment. Some even said they had lost their fear of death.

Armed with these promising results, Griffiths and his colleagues turned their attention to other clinical applications. They decided to investigate tobacco addiction—in part because it is much easier to quantify than emotional or spiritual outcomes. Johns Hopkins researcher Matthew Johnson led a small pilot study in 2014 to see whether psilocybin could help people quit smoking. It was an open-label study, meaning the participants knew they were getting the drug and not a placebo.

The work followed a classic model for psychedelic therapy in which the participant lies on a couch and wears eyeshades while listening to music. Researchers do not talk to or guide subjects during the trip, but before each session, they do try to prepare people for what they might experience. In Johnson and his colleagues' study, participants also underwent several weeks of cognitive-behavioral therapy (talk therapy aimed at changing patterns of thinking) before and after taking psilocybin. The drug was given in up to three sessions—one on the target quit date, another two weeks later and a third, optional one eight weeks afterward. The subjects returned to the lab for the next 10 weeks to have their breath and urine tested for evidence of smoking and came back for follow-up meetings six and 12 months after their target quit date.

At the six-month mark, 80 percent of smokers in the pilot study (12 out of 15) had abstained from cigarettes for at least a week, as verified by Breathalyzer and urine analysis—a vast improvement over other smoking-cessation therapies, whose efficacy rates are typically less than 35 percent. In a follow-up paper, Johnson and his colleagues reported that 67 percent of participants were still abstinent 12 months after their quit date, and 60 percent of them had not smoked after 16 months or more. Additionally, more than 85 percent of the subjects rated their psilocybin trip as one of the five most meaningful and spiritually significant experiences of their lives. The team is currently more than halfway through a larger, five-year study of 80 people randomized to receive either psilocybin or a nicotine patch at the new Johns Hopkins center. Recruitment for the study is ongoing.

The exact brain mechanism by which the therapy appears to work remains unclear. At the psychological level, Johnson says, there is evidence that the sense of unity and mystical significance many people experience on psilocybin is associated with greater success in quitting, and those who take the drug may be better able to deal with cravings. At the biological level, he adds, scientists have hypothesized that psilocybin may alter communication in brain networks, possibly providing more top-down control over the organ's reward system. A team led by Johns Hopkins cognitive neuroscientist Frederick Barrett is now investigating further by using functional magnetic resonance imaging to measure brain activity before and after patients undergo the therapy.

Like any drug, psilocybin comes with risks. People with psychotic disorders such as schizophrenia (or a strong predisposition for them) are generally advised against taking the hallucinogen. People with uncontrolled hypertension are advised to abstain as well, because psilocybin is known to raise blood pressure. Although it appears to be one of the safest "recreational" drugs and is not considered addictive, there have been reports associating it with deaths—but these may have been the result of multiple drugs, impure substances or underlying medical issues. In the smoking study, a third of participants experienced some fear or anxiety at a high dose of the psilocybin, Johnson says. But he adds that the risks can be minimized by carefully selecting participants and administering the drug in a controlled environment.

The smoking study results are promising, but Johnson says its relatively small size is a limitation. Also, subjects in such studies cannot comprise a completely random sample of the population, because it would be unethical to recruit people without telling them they may be taking a psychedelic drug. Thus, participants tend to be people who are open to this category of experience and, potentially, more apt to believe in its efficacy. And it is also hard to tease apart the effects of psilocybin from those of the cognitive-behavioral therapy in the smoking study, Johnson notes. He and his colleagues at the new center plan to conduct a double-blind, placebo-controlled study—the gold standard for medical investigations—in the future. Johns Hopkins researchers are also starting or planning studies using psilocybin therapy for a wide range of other conditions, including opioid addiction, PTSD, anorexia, post-treatment Lyme disease syndrome, Alzheimer's and alcoholism in people with depression.

David Nichols, a professor emeritus of pharmacology at Purdue University, who was not involved in the recent Johns Hopkins studies but had synthesized the psilocybin used in Griffiths's 2006 and 2016 papers, has been conducting research on psychedelics since the late 1960s. Back then, "you probably could have counted on one hand the number of people in the world that were working in this field. There wasn't any money; there was no interest. [Psychedelics] were just looked at as drugs of abuse," he says. Now "there's a whole society set up to study these, with probably 150 international scientists working on it."

Nichols says he has supported Griffiths's and Johnson's work since its early days, as they gathered the initial data that excited wealthy donors enough to fund the latest research. Philanthropic funding "is the way it's going to be—until the National Institutes of Health decide that this is a field worth funding," he says. "There are still too many political considerations that are keeping that from happening, but eventually, we'll get there. We'll get institutional support. We're just not there yet."

Taking Sex Differences in Personality Seriously

New approaches are shedding light on the magnitude of sex differences in personality, and the results are so strong and pervasive that they can no longer be ignored

By Scott Barry Kaufman

JAKE OLIMB *GETTY IMAGES*

Scott Barry Kaufman is a psychologist at Columbia University exploring intelligence, creativity, personality and well-being. In addition to writing the column Beautiful Minds for *Scientific American*, he also hosts The Psychology Podcast, and is author and/or editor of eight books, including *Wired to Create: Unraveling the Mysteries of the Creative Mind* (with Carolyn Gregoire) and *Ungifted: Intelligence Redefined*.

FEW TOPICS IN PSYCHOLOGY ARE MORE CONTROVERSIAL THAN SEX DIFFERENCES[1]. Debates can be classified into two main types: (a) The *description of sex differences,* including both the size and variability of sex differences across a multitude of physical and psychological traits, and (b) the *origins and development of sex differences*, including the complex interplay among social, cultural, genetic and biological factors that influence sex differences.

These lines often get blurred. Researchers who emphasize sociocultural factors in their research tend to conceptualize sex differences as small and worry that if we exaggerate the differences, then all hell will break loose in society. On the other side, those who emphasize biological influences tend to emphasize how differences in personality and behavior can be quite large.

I believe that this blurring between the *descriptive* and the *explanatory* levels of analysis has stunted the field and distorted public debates over these complex and sensitive issues. In order to make real long-lasting changes that actually have an effect on desired outcomes, our knowledge of the truth needs to be as clear as possible.

In this article I will focus on the personality domain, which has made some truly fascinating advances in only the past few years. I will argue that while the science still has a long way to go to fully flesh out the complex interplay of nature and nurture in creating these differences, it's nonetheless time to take sex differences in personality seriously.

MALE AND FEMALE PERSONALITIES

A large number of well-done studies have painted a rather consistent picture of sex differences in personality that are strikingly consistent across cultures (see here, here and here). It turns out that the most pervasive sex differences are seen at the "narrow" level of personality traits, not the "broad" level (see here for a great example of this basic pattern).

At the broad level, we have traits such as extraversion, neuroticism and agreeableness. But when you look at the specific *facets* of each of these broad factors, you realize that there are some traits that males score higher on (on average) and some traits that females score higher on (on average), so the differences cancel each other out. This canceling out gives the appearance that sex differences in personality don't exist when in reality they *very much do exist*.

For instance, males and females *on average* don't differ much on extraversion. At the narrow level, however, you can see that males on average are more assertive (an aspect of extraversion), whereas females *on average* are more sociable and friendly (another aspect of extraversion). So what does the overall picture look like for males and females on average when going deeper than the broad level of personality?

On average, males tend to be more dominant, assertive, risk-prone, thrill-seeking, tough-minded, emotionally stable, utilitarian and open to abstract ideas. Males also tend to score higher on self-estimates of intelligence, even though sex differences in general intelligence measured as an ability are negligible[2]. Men also tend to form larger, competitive groups in which hierarchies tend to be stable and in which individual relationships tend to require little emotional investment. In terms of communication style, males tend to use more assertive speech and are more likely to interrupt people (both men and women) more often—especially intrusive interruptions—which can be interpreted as a form of dominant behavior.

Of course, there are many men who don't display high levels of all these traits. But that fact doesn't contradict the broader pattern. For instance, I can recognize that I am a man who has quite a mix of extremely masculine and extremely feminine personality traits and also recognize that my own personal experience doesn't invalidate the generalizable findings. Which is why I will keep italicizing on average to emphasize that point.

In contrast, females on average tend to be more sociable, sensitive, warm, compassionate, polite, anxious, self-doubting and more open to aesthetics. On average, women are more interested in intimate, cooperative,

dyadic relationships that are more emotion-focused and characterized by unstable hierarchies and strong egalitarian norms. Where aggression does arise, it tends to be more indirect and less openly confrontational. Females also tend to display better communication skills, displaying higher verbal ability and the ability to decode other people's nonverbal behavior. Women also tend to use more affiliative and tentative speech in their language and tend to be more expressive in both their facial expressions and bodily language (although men tend to adopt a more expansive, open posture). On average, women also tend to smile and cry more frequently than men, although these effects are very contextual, and the differences are substantially larger when males and females believe they are being observed than when they believe they are alone.

Contrary to what one might expect, for all these personality effects the sex differences tend to be larger—not smaller—in more individualistic, gender-egalitarian countries. One could make the point that many of these differences aren't huge, and they'd be mostly right if we just stopped our analysis here[3]. But in recent years it's becoming increasingly clear that when you take a look at the overall gestalt of personality—taking into account the correlation between the traits—the differences between the sexes become all the more striking.

THE GESTALT OF PERSONALITY

Personality is multidimensional, which has implications for calculating sex differences in personality. Relatively small differences across multiple traits can add up to substantial differences when considered as a whole profile of traits. Take the human face, for example. If you were to just take a particular feature of the face—such as mouth width, forehead height or eye size—you would have difficulty differentiating between a male face and a female face. You simply can't tell a male eyeball from a female eyeball, for instance. Yet a look at the combination of facial features produces two very distinct clusters of male versus female faces. In fact, observers can correctly determine sex from pictures with greater than 95 percent accuracy[4]. Here's an interesting question: Does the same apply to the domain of personality?

Interestingly, yes. You can calculate a metric called D, which is a summary of how statistically separate two groups are from each other (that is, how good of a line you can draw between groups from a statistical point of view). This metric allows you to take into account how all the personality traits tend to be related to one another in the general population. For instance, people who are conscientious also tend to be more emotionally stable, so if you find someone who is very conscientious and also super neurotic, that person stands out more (has a more unusual personality profile) given the overall correlational structure. With more traits, things get even more interesting. You can have a combination of traits that are less expected, and thus more informative, because they go against the trends of the correlational structure[5].

There now exists four large-scale studies that use this multivariate methodology (see here, here, here and here). All four studies are conducted cross-culturally and report on an analysis of narrow personality traits (which, as you may recall, is where most of the action is when it comes to sex differences). Critically, all four studies converge on the same basic finding: when looking at the overall gestalt of human personality, there is a truly striking difference between the typical male and female personality profiles.

Just how striking? Well, actually, really striking. In one recent study, Tim Kaiser, Marco Del Giudice and Tom Booth analyzed personality data from 31,637 people across a number of English-speaking countries. The size of global sex differences was D = 2.10 (it was D = 2.06 for just the U.S.). To put this number in context, a D = 2.10 means a classification accuracy of 85 percent. In other words, their data suggest that the probability that a randomly picked individual will be correctly classified as male or female based on knowledge of their global personality profile is 85 percent (after correcting for the unreliability of the personality tests).

Consistent with prior research, the researchers found that the following traits are most exaggerated among females when considered separately from the rest of the gestalt: sensitivity, tender-mindedness, warmth, anxiety, appreciation of beauty and openness to change. For males, the most exaggerated traits were emotional stability, assertiveness/dominance, dutifulness, conservatism, and conformity to social hierarchy and traditional structure.

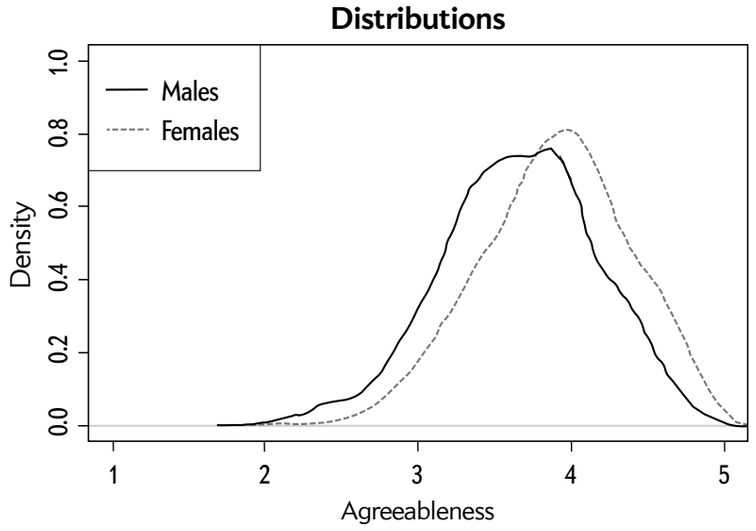
This basic pattern of findings was replicated in another recent large-scale survey of narrow personality traits conducted on nearly a million people across 50 countries. Using different personality tests and averaging across all countries, Kaiser found a D = 2.16, which is very similar to the effect size found in the other study on English-speaking countries. While there was cross-cultural variation in the effect, there was a general trend for more developed, individualistic countries with higher food availability, less pathogen prevalence and higher gender equality to show the largest sex differences in global personality[6].

In particular, Scandinavian countries consistently showed larger-than-average sex differences in global personality, together with the U.S., Canada, Australia, the U.K. and other Northern and Eastern European countries. The countries with the smallest sex differences in global personality included several Southeast Asian countries. To be sure, there wasn't a perfect correlation between more developed, gender-egalitarian countries and sex differences (for example, Russia displayed the largest sex difference with D = 2.48). But even Pakistan—the country with the smallest sex differences in global personality in the world, according to this study—had a D = 1.49. This means that even when you look around the world

for the country with the smallest sex difference in global personality, the classification accuracy of that country is still 77 percent!

These numbers dovetail with a number of studies showing a similar level of classification looking at whole brain data. By applying a multivariate analysis of the whole brain, researchers are now able to classify whether a brain is male or female with 77 to 93 percent accuracy (see here, here, here, here and here). In fact, some recent studies using the most sophisticated techniques have consistently found greater than 90 percent accuracy rates looking at whole brain data (see here, here and here). While this level of prediction is definitely not perfect—and by no means do those findings justify individual stereotyping or discrimination—that's really high accuracy as far as science goes[7].

All these data are really hard to ignore and dismiss out of hand. But what are the implications?



Overlapping distributions of agreeableness for men and women. Vertical axis indicates density, or the proportion of the sample in a given area under the curve.

IMPLICATIONS

All the findings I've presented up to this point are merely descriptive; they don't prescribe any particular course of action, and they do not say anything about the complex interplay of genetic and cultural influences that may cause these differences to arise in the first place. It is very difficult finding evidence that would indicate just how much of sex differences stem from society versus genetics (although it's most certainly a mix; more on that later). Even the brain findings discussed above don't reveal the causes of brain development. Experience is constantly sculpting brain development.

But even if we just stay at the descriptive level, there are still a number of very important implications of the existence of large sex differences in personality. For one, the multivariate findings may help answer a question people have been puzzling about in psychology for quite a while: Why do we have all these studies showing that male and female behaviors are so similar, yet people in everyday life continue to think as if males and females were very separable? It is possible that people in everyday life are actually closer to the truth because when we reason about personality, we rarely reason about one trait at a time.

If people do indeed create a gestalt in personality perception, then the relevant analysis is a multivariate analysis, not a univariate analysis (which has been the predominant method in the field for so long). "People might be more reasonable than you think," Del Giudice, a leader in the science of sex differences, told me, "Why would you expect people to just make up differences between men and women that aren't there? One possibility is that they are not making it up. What they are considering when they are thinking about men and women is not just one trait at a time, but a combination of traits."

Another possible factor that may help further our understanding of pervasive stereotypical expectations may also have to do with recognizing the importance of the tails of the personality distribution. Even relatively small differences at the average level can lead to very large differences in the proportion of groups at the extremes. For instance, if you look at the density distribution for agreeableness, the average difference between males and females is only about 0.4 of a standard deviation. But if you look closely, you can see that there are way more women than men who are super agreeable and way more men than women who are super disagreeable. It's likely that the behaviors carried out by those tails have a huge impact on society—on social media, in politics, in the boardroom and even in the bedroom.

Now, one might counter at this point: Scott, you really should stop talking openly and honestly about these findings and implications because if the truth got out there, it could cause harm. But here's the thing: rarely do we consider the harm that could be caused by ignoring sex differences! One can think of many ways in which pretending something doesn't exist may actually cause greater harm psychologically than accepting the facts of the matter. As Del Giudice put it to me:

> "People don't want to just give up on trying to understand the world. They want to make sense of the world. And so, if the right explanation is that there is some kind of difference, and you kind of close off that possible explanation because of ideological reasons, it's not like people stop asking why. They will come up with a different explanation. So you will get a chain of worse and worse and worse explanations that may actually backfire in all sorts of ways."

Take heterosexual marriage. Many couples go into a marriage assuming that sex differences in personality are minimal. We know that on average, however, females in relationships want constant emotional connections, whereas on average men don't tend to be equally as inter-

ested in that aspect of the relationship. An incredible amount of stress in a marriage may result from what people are expecting about each other based on the assumption that everything has to be equal and both partners must feel the same exact way about everything. But here's the thing: we don't all have to be the same in every dimension in order to appreciate and respect each other.

Of course, couples need to work out the fit between their very special and unique personalities. I am a strong believer that individual differences are more important than sex differences. Nevertheless, sex differences are also part of the picture and may be particularly detrimental to a relationship if all partners go into the marriage thinking that they "should not exist," instead of coming to a healthy acceptance of sex differences, even laughing about them and attempting to understand differences in interests and motivations that fall along sex-related lines. Of course, there will be so many aspects of overlap among males and females in a relationship, but there may be a few meaningful differences that on average could be truly impactful and explanatory in predicting relationship satisfaction and understanding.

TOWARD A MATURE, NUANCED AND SOPHISTICATED SCIENCE OF SEX DIFFERENCES

I believe it's time for a more mature, honest and nuanced public discourse about these obviously sensitive yet incredibly important issues.

First and foremost, I think this requires a recognition that none of the findings I presented in this article nor any findings that will ever come out—justifies individual discrimination. We should treat all people as unique individuals first and foremost. No matter what the science says, if an individual shows the interest and ability to enter a field in which their sex is extremely underrepresented (for example, women in math and science or men in nursing and education), we should absolutely be encouraging that individual to enter the field and do everything we can to help them feel a sense of belonging. I may be weird, but I don't see any contradiction whatsoever between being an advocate for equitable opportunity for all people and being an equally strong advocate for respecting scientific findings and attempting to get as close as possible to the truth about average sex differences.

I also believe that a truly mature, honest and nuanced discussion of the origins of sex differences must recognize the deep influence of genetics and biology[8]. That doesn't mean that we ignore sociocultural factors, which are clearly important. But sex differences in behavior are so pervasive in nearly every other species. It's just not plausible that somehow male and female psychology evolved to be identical despite the physiological differences and different reproductive roles across human evolutionary history.

This is why biologically oriented folks draw on a wide range of explanatory concepts from biology, as well as cross-cultural, anthropological and primatological evidence about present-day and ancient humans and their primate relatives. This doesn't mean that such theories are always right. The point is that the methodology is far richer and systematic than they are so often treated in the popular media. The best sources to counteract this misconceptions are Dave Geary's book *Male, Female* and Steve Stewart-Williams's *The Ape That Understood the Universe*. If you want to dive into a more academic treatise, consult this academic paper by John Archer.

I'm actually really optimistic that such discussions don't have to devolve into polarization and ad hominem name calling, with accusations of "sexism" on one side and being "antiscience" on the other side. I'm optimistic because I think a great example of a mature debate on the this topic already exists.

In February 2019 psychologists Cordelia Fine, Daphna Joel and Gina Rippon wrote an article called "Eight Things You Need to Know about Sex, Gender, Brains, and Behavior: A Guide for Academics, Journalists, Parents, Gender Diversity Advocates, Social Justice Warriors, Tweeters, Facebookers, and Everyone Else." Based on their many years observing both the scientific and popular treatment of the topic of sex differences in brain and behavior, the authors provide an accessible guide to help everyone interpret new biological findings. They rightly point out that people unfortunately tend to unthinkingly ascribe the mere existence of sex differences to "immutable biological factors," an assumption that does not automatically follow from the data. Not only that, but it's true that there is very little biologically that's "immutable" other than the genetic sequence, a fact that is widely known among all the psychologists that I know.

Del Guidice, Geary, David Puts and David Schmitt then wrote eight counterpoints to their article, agreeing with some of their premises but disagreeing with other premises. They argue that Fine and her colleagues assume that most sex differences are small, inconsistent, highly malleable and for the most part socially constructed and argue that:

> "minimizing the magnitude of important sex differences and discounting their biological origins can be just as damaging (for science and society at large) as exaggerating them and accepting simplistic biological explanations of sex differences at face value.... An honest, sophisticated public debate on sex differences demands a broad perspective with an appreciation for nuance and full engagement with all sides of the question."

In a response to their counterpoint, Fine, Joel and Rippon note their pleasure at Del Giudice and his colleagues'

response but point out several places of "ghost disagreement"—that is, places where Del Giudice and his colleagues argued against views that they did not express and actually do not hold.

This back and forth was such a great example of the importance of constructive debate and giving people enough benefit of the doubt to allow them to clarify their views so that they aren't misinterpreted or their views aren't taken out of proportion. Fine and her colleagues concluded that "exchanges such as the present one, when focused on evidence and claims, are valuable—and rarer than we would like." For anyone who wants to dive deeper into these complex debates and see a great example of how real progress can be made in furthering knowledge and understanding, I highly recommend reading this entire exchange.

In my view, a more mature, sophisticated and nuanced understanding of sex differences in personality and behavior is possible. One important step is to take sex differences in personality seriously. Only by facing reality as clearly as possible can we even begin to make changes that will have a real positive impact on everyone.

END NOTES

[1]Because of the research that has already been conducted on this topic, I intentionally used the phrase "sex" differences in this article rather than "gender" differences—sex defined as a collection of traits (for example, X/Y chromosomes, gonads, hormones and genitals) that cluster together in about 99.98 percent of humans (see here and here). Of course, I do not mean to suggest that the exceptions to the sex binary are unimportant, and I fully believe that all variations in gender identity and sexual orientation are amenable to scientific investigation and deserve to be studied in their full richness. Also, I think it's an interesting and open question the extent to which there are gender differences in personality, especially among the many different gender identities that people are adopting in recent years. I'd definitely be interested in seeing more research looking into that question as well.

[2]It should be noted, however, that men are typically found to show more variance in general cognitive ability scores than women (see here and here).

[3]One notable exception is an interest in people versus an interest in things. The sex differences on this dimension are actually quite large, with some large studies finding greater than one standard deviation of a difference between males and females on average on this dimension (see here and here).

[4]I could see someone being concerned that this finding somehow strips us of our individuality—that essence of us that transcends our biological sex. Yet I think that fear is unwarranted. After all, there now exist really sophisticated apps in which you can change the sex of your face, but even then, you still remain recognizable. I think maintaining one's individuality doesn't contradict the generalizable findings regarding the high classification rates of sex based on one's physical characteristics.

[5]To be sure, the multivariate approach (where you look at personality as a whole) isn't always better than a more univariate approach (where you focus on a specific variable). It's all about context and what you are trying to predict and your purposes of prediction. For instance, if what you are trying to predict is clearly based on a particular subset of traits, then just adding more traits into the model may produce an illusory effect. There are a few criticisms of the multivariate approach, however, that really do not hold water (see here). One is the criticism that a multivariate approach to personality doesn't say anything meaningful, because it's not valid to aggregate traits in a multivariate analysis. This is a fair criticism for domains that include a hodgepodge of traits that don't go together in any meaningful way. But that doesn't apply to the domain of personality. There exists a plethora of research across cultures on the correlational structure of personality. Of course, if you start adding irrelevant variables such as shoe size, voting preference or height to the personality data you will get an artificially big separation between the sexes, and it wouldn't tell us much of anything meaningful. That's not how these studies are conducted, however. A second potential criticism is that the more traits you throw into a multivariate analysis, of course the effects are going to get bigger and bigger and bigger. So it's not interesting that we get these big effects. While this criticism is true—technically speaking, the more traits you add, the more differences will grow and will never shrink—it's simply not true that the differences will keep growing at the same rate. Because the multivariate analysis takes into account the correlation between the traits, you will eventually start seeing less of an effect of adding in additional personality traits because additional traits will start becoming more and more redundant.

[6]Interestingly, Kaiser found that after controlling for some potential confounds relating to ecological stress, only historic pathogen prevalence, food availability and cultural individualism were still correlated with sex differences in personality (the specific correlation between the gender equality of the country and sex differences was reduced to zero after controlling for confounds). Kaiser concludes that "[previously] reported correlations between greater sex differences and outcomes of gender equality could be due to confounding by influences of ecological stress."

[7]Someone may look at these studies and say: Well, what about this *New York Times* op-ed: "Can We Finally Stop Talking about 'Male' and 'Female' Brains?" It turns out that the data that are mentioned in that study conducted by Daphna Joel and her colleagues (see here) were not based on whole brain data. This matters. The researchers left it to the reader to infer that their findings also apply to whole brains by extension, but it turns out that such an

extension is not warranted given the recent spate of studies that are all converging on 77 to 93 percent classification accuracy based on whole brain data—including a more recent study led by Joel! What's more, the method that Joel and her colleagues devised for quantifying "internal consistency" in their earlier article is a straw man guaranteed to always find very low levels of consistency. By defining "consistency" as 100 percent uniformity, there is no way that their method will ever detect consistency as long as there is some variation within each sex. Del Giudice and his colleagues have shown this to be the case with artificial data and illustrated it by showing that the method cannot even detect consistency within species (they compared the facial anatomy of different species of monkeys). More realistic than having 100 percent consistency, in my view, is whether the pattern is statistically robust—whether you can distinguish between men and women with a very high degree of accuracy based on aggregate patterns of interests. And this is why their initial finding is such a red herring: their conclusion is not based on whole brain data. To dive deeper into the critique of the study by Joel and her colleagues, I recommend reading this and this.

[8]I intentionally separated out "genetic" from "biological" in this sentence because it's a common misconception that "biological" equates to "genetic." The question "Are sex differences biological or cultural?" is actually a meaningless question because every sex difference is biological when it's expressed, regardless of whether its origins are cultural or genetic. Social-learning processes are biological. Aspects of personality that are learned are also biological. In fact, anything that affects behavior is acting biologically on the brain. When people say traits or sex differences are "biological," they probably really mean "genetic."

What's Next for Psychology's Embattled Field of Social Priming

A promising field of research on social behavior struggled after investigators couldn't repeat key findings. Now researchers are trying to establish what's worth saving

By Tom Chivers

KAROL BANACH

THREE YEARS AGO a team of psychologists challenged 180 students with a spatial puzzle. The students could ask for a hint if they got stuck. But before the test, the researchers introduced some subtle interventions to see whether these would have any effect.

The psychologists split the volunteers into three groups, each of which had to unscramble some words before doing the puzzle. One group was the control, another sat next to a pile of play money and the third was shown scrambled sentences that contained words relating to money.

The study, published last June, was a careful repeat of a widely cited 2006 experiment. The original had found that merely giving students subtle reminders of money made them work harder: in this case, they spent longer on the puzzle before asking for help. That work was one among scores of laboratory studies that argued that tiny subconscious cues can have drastic effects on our behavior.

Known by the loosely defined terms "social priming" or "behavioral priming," these studies include reports that people primed with "money" are more selfish; that those primed with words related to professors do better on quizzes; and even that people exposed to something that literally smells fishy are more likely to be suspicious of others.

The most recent replication effort, however, led by psychologist Doug Rohrer of the University of South Florida, found that students primed with "money" behave no differently on the puzzle task from the controls. It is one of dozens of failures to verify earlier social-priming findings. Many researchers say they now see social priming not so much as a way to sway people's unconscious behavior but as an object lesson in how shaky statistical methods fooled scientists into publishing irreproducible results.

This is not the only area of research to be dented by science's "replication crisis." Failed replication attempts have cast doubt on findings in areas from cancer biology to economics. But so many findings in social priming have been disputed that some say the field is close to being entirely discredited. "I don't know a replicable finding. It's not that there isn't one, but I can't name it," says Brian Nosek, a psychologist at the University of Virginia, who has led big replication studies. "I've gone from full believer to full skeptic," adds Michael Inzlicht, a psychologist at the University of Toront and an associate editor at the journal *Psychological Science*.

Some psychologists say the pendulum has swung too far against social priming. Among these are veterans of the field who insist that their findings remain valid. Others accept that many of the earlier studies are in doubt but say there's still value in social priming's central idea. It is worth studying whether it's possible to affect people's behavior using subtle, low-cost interventions—as long as the more outlandish and unsupported claims can be weeded out, says Esther Papies, a psychologist at the University of Glasgow in Scotland.

Equipped with more rigorous statistical methods, researchers are finding that social-priming effects do exist but seem to vary between people and are smaller than first thought, Papies says. She and others think that social priming might survive as a set of more modest, yet more rigorous, findings. "I'm quite optimistic about the field," she says.

RISE AND FALL

The roots of the priming phenomenon go back to the 1970s, when psychologists showed that people get faster at recognizing and processing words if they are primed by related ones. For instance, after seeing the word "doctor," they recognized "nurse" faster than they did unrelated words. This "semantic" priming is now well established.

But in the 1980s and 1990s researchers argued that priming could affect attitudes and behaviors. Priming individuals with words related to "hostility" made them more likely to judge the actions of a character in a story as hostile, a 1979 study found. And in 1996 John

Tom Chivers is a science journalist based in London.

Bargh, a psychologist at New York University, found that people primed with words conventionally related to age in the U.S.—"bingo," "wrinkle," "Florida"—walked more slowly than the control group as they left the lab, as if they were older.

Dozens more studies followed, finding that priming could affect how people performed at general-knowledge quizzes, how generous they were or how hard they worked at tasks. These behavioral examples became known as social priming, although the term is disputed because there is nothing obviously social about many of them. Others prefer "behavioral priming" or "automatic behavior priming."

In his 2011 best-seller *Thinking, Fast and Slow,* Nobel Prize–winning psychologist Daniel Kahneman mentioned several of the best-known priming studies. "Disbelief is not an option," he wrote of them. "The results are not made up, nor are they statistical flukes. You have no choice but to accept that the major conclusions of these studies are true."

But concerns were starting to surface. That same year Daryl Bem, a social psychologist at Cornell University, published a study suggesting that students could predict the future. Bem's analyzing relied on statistical techniques that psychologists regularly used. "I remember reading it and thinking, 'If we can do this, we have a problem,'" says Hans IJzerman, a social psychologist at the University of Grenoble Alps in France.

Also that year three other researchers published a deliberately absurd finding: that those who listened to the Beatles song "When I'm Sixty-Four" literally became younger than a control group that listened to a different song. They achieved this result by analyzing their data in many different ways, getting a statistically significant result in one of them by simple fluke and then not reporting the other attempts. Such practices, they said, were common in psychology and allowed researchers to find whatever they wanted, given some noisy data and small sample sizes.

The papers had an explosive impact. Replication efforts that cast doubt on key findings started to appear, including a 2012 report that repeated Bargh's aging study and found no effect of priming unless the people observing the experiment were told what to expect. It did not help that this all took place as it was discovered that a leading social psychologist in the Netherlands, Diederik Stapel, had been faking data for years.

In 2012 Kahneman wrote an open letter to Bargh and other "students of social priming," warning that "a train wreck" was approaching. Despite his being a "general believer" in the research, Kahneman worried that fraud such as Stapel's, replication failures and a tendency for negative results not to get published had created "a storm of doubt."

Eight years later the storm has uprooted many of social priming's flagship findings. Eric-Jan Wagenmakers, a psychologist at the University of Amsterdam, says that when he read the relevant part of Kahneman's book, "I was like, 'not one of these studies will replicate.' And so far nothing has."

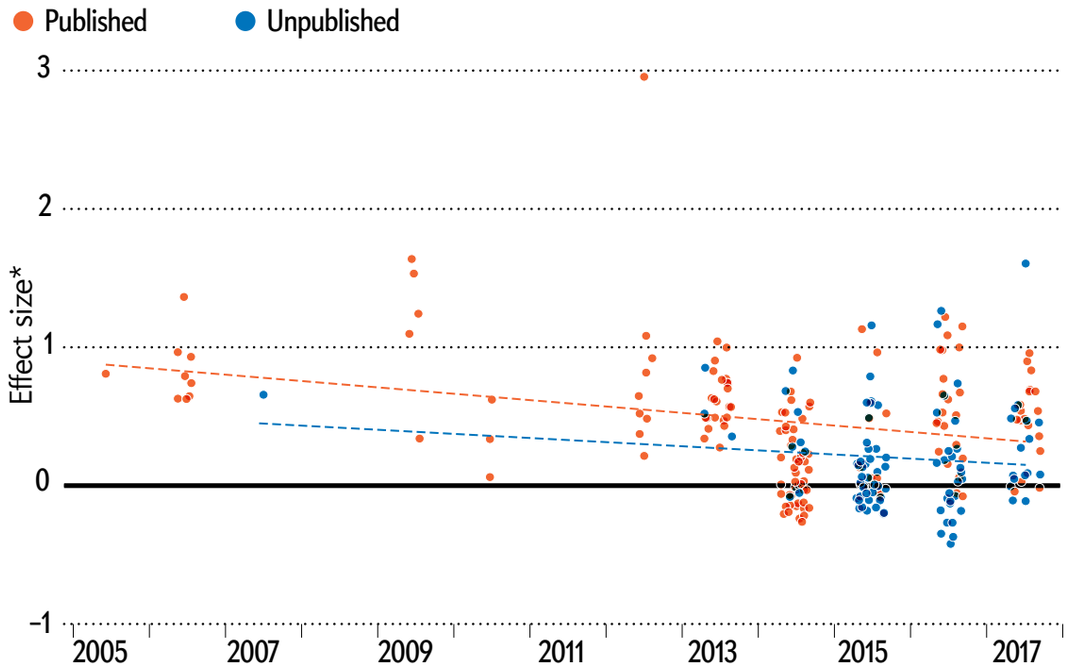
Psychologist Eugene Caruso reported in 2013 that reminding people of the concept of money made them more likely to endorse free-market capitalism. Now at the University of California, Los Angeles, Caruso says that having tried bigger and more systematic tests of the effects, "there does not seem to be robust support for them." Ap Dijksterhuis, a researcher at Radboud University in Nijmegen, the Netherlands, says that his paper suggesting that students primed with the word "professor" do better at quizzes "did not pass the test of time."

Kahneman told *Nature:* "I am not up-to-date on the most recent developments, so should not comment."

Researchers had been whispering about not being able to repeat big findings years before the priming bubble began to burst, Nosek says. Afterward, in lessons shared with science's wider replication crisis, it became clear that many of the problematic findings were proba-

Waning Effect

A meta-analysis of 246 experiments that exposed people to money-related stimuli found that early studies reported larger priming effects on behavior, emotions and attitudes than did later ones. It also revealed larger effects in published work than in unpublished experiments provided by authors of the original studies.



*Effect size measured by a value known as Hedges' g, where 1 indicates that primed and control groups differed by 1 standard deviation.

bly statistical noise—fluke results garnered from studies on too-small groups of people—rather than the result of fraud. It seems that many researchers were not alert to how easy it is to find significant-looking but spurious results in noisy data. This is especially so if researchers "HARK" (Hypothesize After Results are Known)—that is, change their hypotheses after looking at their data. The fact that journals tend not to publish null results didn't help, because it meant the only findings that got through were the surprising ones.

There is also evidence that subconscious experimenter effects have been a problem, Papies says: one study found that when experimenters were aware of the priming effect they were looking for, they were much more likely to find it, suggesting that, subconsciously, they would affect the results in some way.

Since then, there have been widespread moves throughout psychology to improve research methods. These include preregistering study methods before looking at data, which prevents HARKing, and working with larger groups of volunteers. Nosek, for instance, has led the Many Labs project, in which undergraduates at dozens of labs try to replicate the same psychology studies, giving sample sizes of thousands. On average, about half of the papers that Many Labs looks at can be replicated successfully. Other collaborative efforts include the Psychological Science Accelerator, a network of labs that work together to replicate influential studies.

"If preregistration stops people from HARKing, then I guess it's good. But it always struck me as an insult. 'We don't trust you to be honest'; it feels like we're being treated like criminals, wearing ankle bracelets."

—John Bargh

THE NEW SOCIAL PRIMING

Today much of the work being done in social priming involves replications of earlier work or meta-analyses of multiple papers to try to tease out what still holds true. A meta-analysis of hundreds of studies on many kinds of money priming, reported last April, found little evidence for the large effects the early studies claimed. It also found larger effects in published studies than in unpublished experiments that had been shared with the authors of the meta-analysis.

Original work hasn't dried up entirely, Papies says, although the focus is changing. Much of the high-profile social-priming work of the past was designed to find huge, universal effects, she says. Instead her group's studies focus on finding smaller effects in the subset of people who already care about the thing being primed. She has found that people who want to become thinner are more likely to make healthy food choices if they are primed, say, with words on a menu such as "diet," "thin" and "trim figure." But it works only in people for whom a healthy diet is a central goal; it doesn't make everyone avoid fattening foods.

This matches the findings of a meta-analysis from 2015, led by psychologist Dolores Albarracín of the University of Illinois at Urbana-Champaign. It looked at 352 priming studies that involved presenting words to people, and it found evidence of real, if small, effects when the prime was related to a goal that the participants cared about. That analysis, however, deliberately looked only at experiments in which the priming words were directly related to the claimed effect, such as rudeness-related words leading to ruder behavior or attitudes. It avoided looking at studies with primes that had what it termed "metaphorical" meaning—including the aging-related words Bargh said led to slower walking or the money-related priming work.

Research into priming has declined, however, and what is considered priming is not always the same as the startling claims of the 1990s and 2000s. "There's a lot less than there was five or 10 years ago," says Antonia Hamilton, a neuroscientist at University College London, who still works on priming. Partly, she says, that's because of the replication problems: "We do less since it all blew up. It's harder to make people believe it, and there are other topics that are easier to study." It might also be simply that the topic has become less fashionable, she says.

Hamilton's own work involves, among other things, putting people in functional magnetic resonance imaging (fMRI) scanners to see how priming affects brain activity. In one 2015 study, she used a scrambled-sentence task to prime prosocial ideas (such as helping) and antisocial ones (such as annoying) to see whether it made participants quicker to mimic other people's actions and whether there were detectable differences in brain scans.

Using fMRI is only practical with small numbers of volunteers, so she looks at how the same people respond when they have been primed and when they haven't: a within-subjects design, in contrast to the between-subjects design of priming studies that use a control group. The design means that researchers don't have to worry about preexisting differences between groups, Hamilton says. Her research has found priming effects: people primed with prosocial concepts behave in more proso-

“I still have no doubts whatsoever that in real life, behavior priming works, despite the fact that in the old days, we didn’t study it properly relative to current standards.”
—Ap Dijksterhuis

cial ways, and fMRI scans did show differences in activity in brain areas such as the medial prefrontal cortex, which is involved in regulating social behaviors. But, she says, the effects are more modest than those the classic priming studies found.

Some researchers say that however efforts to test older results pan out, the concept of social or behavioral priming still has merit. “I still have no doubts whatsoever that in real life, behavior priming works, despite the fact that in the old days, we didn’t study it properly relative to current standards,” Dijksterhuis says.

Bargh notes that despite many researchers now discounting them, important early advances do exist—such as his own 2008 study, which reported that holding warm coffee made people behave more warmly toward others. Direct replications have failed to support the result, but Bargh says that a link between physical warmth and social warmth has been demonstrated in other work, including neuroimaging studies.

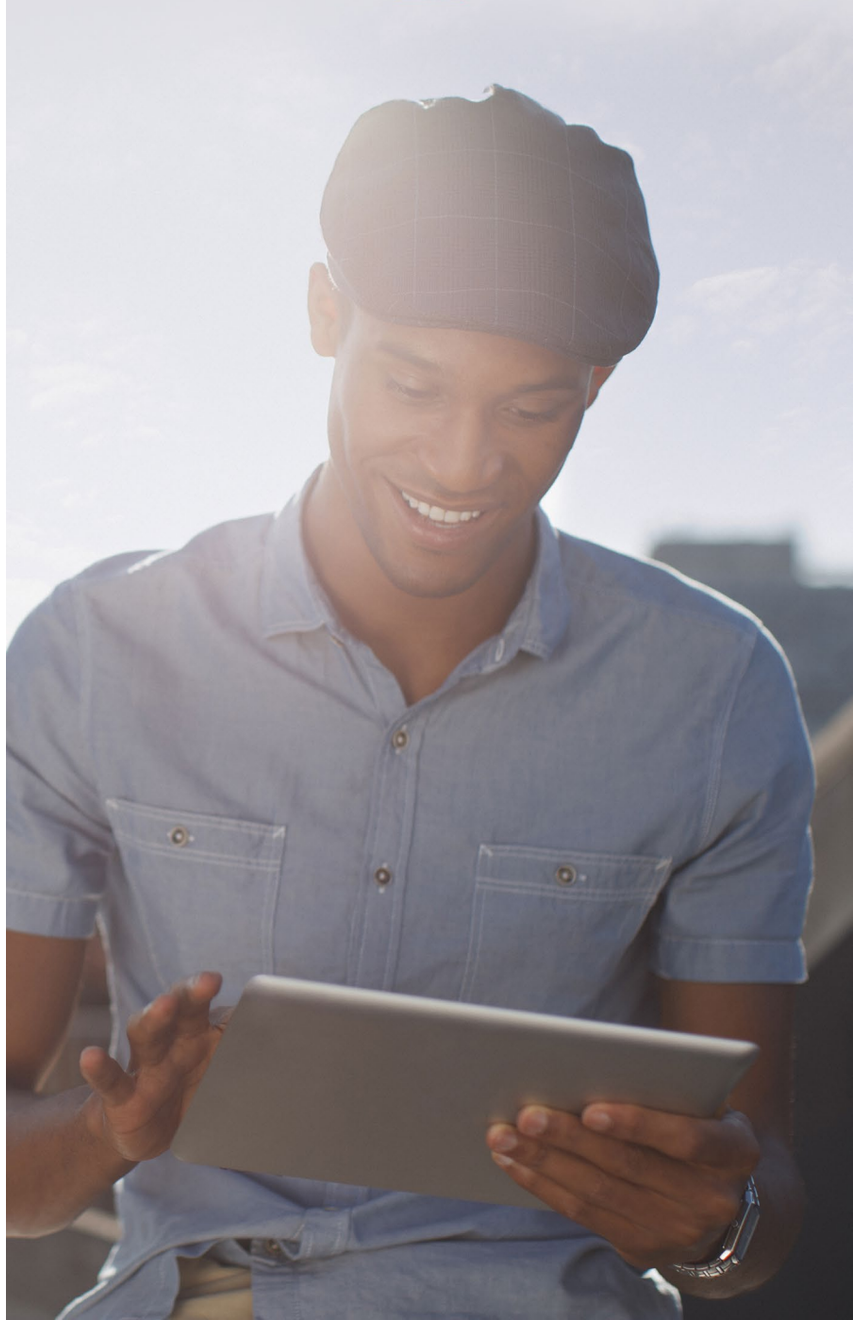
“People say we should just throw out all the work before 2010, the work of people my age and older,” Bargh says, “and I don’t see how that’s justified.” He and Norbert Schwarz, a psychologist at the University of Southern California, say that there have been replications of their earlier social-priming results—although critics counter that these were not direct replications but “conceptual” ones, in which researchers test a concept using related experimental set-ups.

Bargh says that results of social priming are still widely believed and used by nonacademics, such as political campaigners and business marketers, even when they are skeptical. Gary Latham, for instance, an organizational psychologist at the University of Toronto, says: “I strongly disliked Bargh’s findings and wanted to show it doesn’t work.” Despite this, he says, he has for 10 years consistently found that priming phone marketers with words related to ideas of success and winning increases the amount of money they make. But Leif Nelson, a psychologist at the University of California, Berkeley, emphasizes that whether or not social-priming ideas are subsequently confirmed, the classic studies in the field were not statistically powerful enough to detect the things they claimed to find.

Bargh sees positives and negatives in how psychology research has changed. “If preregistration stops people from HARKing, then I guess it’s good,” he says, “but it always struck me as an insult. “We don’t trust you to be honest’; it feels like we’re being treated like criminals, wearing ankle bracelets.”

Others disagree. The move toward open, reproducible science, according to most psychologists, has been a huge success. Social priming as a field might survive, but if it does not, then at least its high-profile problems have been crucial in forcing psychology to clean up its act. “I have to say I am pleasantly surprised by how far the field has come in eight years,” Wagenmaker says. “It’s been a complete change in how people do things and interpret things.”

This article is reproduced with permission and was first published in Nature *on December 11, 2019.*



Brett Beasley is associate director of the Notre Dame Deloitte Center for Ethical Leadership and a term assistant teaching professor of management and organization at the Mendoza College of Business at the University of Notre Dame.

OBSERVATIONS

How Disinformation Hacks Your Brain

The digital age has heightened our vulnerability to falsehood, but recognizing such weaknesses can help guard against them

GETTY IMAGES

Three years ago Edgar Welch sent a text message to a friend announcing he was "Raiding a pedo ring, possibly sacraficing [sic] the lives of a few for the lives of many." Two days later, he drove 350 miles to a Washington, D.C., pizza parlor called Comet Ping Pong and entered with a .38 revolver and an AR-15 semiautomatic rifle. He fired shots inside in an attempt to investigate what he believed was a child sex ring with ties to top Democratic Party leaders and sent restaurant patrons and staff fleeing in fear. The sex ring was fake news. The consequences, however, were real. Welch left the premises under arrest and later pled guilty to local and federal weapons charges.

At the time of Welch's disinformation-driven rampage, "post-truth" had just recently entered the public imagination. A few weeks before Welch's arrest, Oxford Dictionaries declared it the word of the year. Many people still struggled to understand how a polite, soft-spoken person like Welch could be led so far from reality. But as the disinformation age has continued to develop over the past three years, science has not stood still. It has given us a more detailed picture than ever of the ways that disinformation hacks our truth judgments.

If the picture is detailed, it is also disconcerting. It suggests that you and I are probably not so different from Welch as we might like to think. Take for example, what happens when we are subjected to repeated false claims. In a recent study, a research team led by Jonas De keersmaecker found that even those of us who are in-

telligent, analytical and comfortable with ambiguity find statements more believable simply because we have heard them repeated.

This phenomenon, known as the illusory truth effect, was first documented in the 1970s, but it is more relevant than ever in the era of fake news. One might immediately think of Donald Trump, who is a prolific peddler of this type of untruth. The *Washington Post* recently reported that there are "more than 350 instances in which [Trump] has repeated a variation of the same claim at least three times." In fact, Trump has repeated some false claims more than 200 times—for example, his claim that his border wall is being built. Of course, there's nothing new about this type of huckster's grift. But online environments supercharge it. They give repeated false claims instant global distribution. More important, they allow the person making false claims to go on doing so while dodging the pressure (and potential legal repercussions) that accompany similar claims in public or in traditional news sources.

Psychologists say that what makes repeated claims seem truer is their "fluency." Fluency means the cognitive ease with which we process a claim. Repeated claims are easier to represent and comprehend. For that reason, they just feel good. Our minds take this feeling as a cue that the claim is true.

In a recent review of the research, Nadia M. Brashier and Elizabeth J. Marsh identify two additional ways disinformation hacks our truth judgments. One that is closely related to fluency and the good feelings it generates is memory. The information and experiences stored in our memory are powerful weapons in the fight for truth. But, as with fluency, we take our memories as cues, not as the raw materials for forming well-considered judgments. We tend, in other words, to go with "good enough." We often accept claims as true when they only partially fit with what we know or remember.

Additionally, we can fall prey to the illusion of explanatory depth, a tendency to overestimate our knowledge and understanding of the issues we care about. Research shows that when we do, we are more likely to hold extreme beliefs and to accept fake news as true.

Unfortunately, digital tools may be making our memories even weaker and less effective for judging truth. As Brashier and Marsh point out, "search algorithms return content based on keywords, not truth. If you search 'flat Earth,' for example, Google dutifully returns photoshopped pictures for a 150-foot wall of ice that keeps us from slipping off the planet." For this reason, relying on the Internet as truth-on-demand rather than looking to our memories and acquired knowledge can backfire in serious ways.

Brashier and Marsh also point out a more basic mismatch between our brains and the digital environment: We tend to make truth our default judgment. This is especially true for visual information. As with the other cues we use to form truth judgments, this is a handy and useful adaptation in other contexts. After all, humans lived for millennia in an environment where we could trust most of our senses most of the time. Now, however, we find ourselves in a new information ecosystem, one in which, according to some sources, we will soon consume more false media than true media. When it comes to coping with that magnitude of misinformation, our brains are simply not well equipped.

Is there anything we can we do to keep our guard up in the post-truth era? We know that simply fact-checking claims is not enough. After all, Welch's "pedo ring" conspiracy theory had been debunked long before he showed up armed at Comet Ping Pong's door.

There are, however, causes for hope. Once we recognize our vulnerabilities, we can recognize many other ways to design our information consumption with them in mind. Along with Emmaline Drew Eliseev, Brashier and Marsh found they could wipe out the illusory truth effect by simply prompting study participates to behave like fact checkers.

One of the most interesting solutions may be a collaborative one. Ziv Epstein, Gordon Pennycook and David G. Rand have found that crowdsourced judgments about the trustworthiness of news sources can be surprisingly accurate. They suggest allowing users of social media to train algorithms to spot fake news as a scalable, decentralized solution. After ignoring warnings from friends and trying unsuccessfully to recruit them, Edgar Welch went it alone. Perhaps if we come together to protect against the vulnerabilities we all share, no one else will make the same mistake.

Viorica Marian is Ralph and Jean Sundin Endowed Professor of communication sciences and disorders and professor of psychology at Northwestern University. Follow her on Twitter @VioricaMarian1

OBSERVATIONS

The Language You Speak Influences Where Your Attention Goes

It's all because of the similarities between words

GETTY IMAGES

Psycholinguistics is a field at the intersection of psychology and linguistics, and one if its recent discoveries is that the languages we speak influence our eye movements. For example, English speakers who hear *candle* often look at a candy because the two words share their first syllable. Research with speakers of different languages revealed that bilingual speakers not only look at words that share sounds in one language but also at words that share sounds across their two languages. When Russian-English bilinguals hear the English word *marker,* they also look at a stamp, because the Russian word for stamp is *marka*.

Even more stunning, speakers of different languages differ in their patterns of eye movements when no language is used at all. In a simple visual search task in which people had to find a previously seen object among other objects, their eyes moved differently depending on what languages they knew. For example, when looking for a clock, English speakers also looked at a cloud. Spanish speakers, on the other hand, when looking for the same clock, looked at a present because the Spanish names for clock and present—*reloj* and *regalo*—overlap at their onset.

The story doesn't end there. Not only do the words we hear activate other, similar-sounding words—and not only do we look at objects whose names share sounds or letters even when no language is heard—but the translations of those names in other languages become activated as well in speakers of more than one language. For example, when Spanish-English bilinguals hear the word *duck* in English, they also look at a shovel because the translations of *duck* and *shovel*—*pato* and *pala,* respectively—overlap in Spanish.

Because of the way our brain organizes and processes linguistic and nonlinguistic information,

a single word can set off a domino effect that cascades throughout the cognitive system. And this interactivity and co-activation is not limited to spoken languages. Bilinguals of spoken and signed languages show co-activation as well. For example, bilinguals who know American Sign Language and English look at cheese when they hear the English word *paper* because *cheese* and *paper* share three of the four sign components in ASL (hand shape, location and orientation but not motion).

What do findings like these tell us? Not only is the language system thoroughly interactive with a high degree of co-activation across words and concepts, but it also impacts our processing in other domains such as vision, attention and cognitive control. As we go about our everyday lives, how our eyes move, what we look at and what we pay attention to are influenced in direct and measurable ways by the languages we speak.

The implications of these findings for applied settings range from consumer behavior (what we look at in a store) to the military (visual search in complex scenes) and art (what our eyes are drawn to). In other words, it is safe to say that the language you speak influences how you see the world not only figuratively but also quite literally, down to the mechanics of your eye movements.

Steve Taylor is a senior lecturer in psychology at Leeds Beckett University in England. He is author of *Spiritual Science: Why Science Needs Spirituality to Make Sense of the World*.

OBSERVATIONS

How a Flawed Experiment "Proved" That Free Will Doesn't Exist

It did no such thing—but the result has become conventional wisdom nonetheless

GETTY IMAGES

In the second half of the 19th century, scientific discoveries—in particular, Charles Darwin's theory of evolution—meant that Christian beliefs were no longer feasible as a way of explaining the world. The authority of the Bible as an explanatory text was fatally damaged. The new findings of science could be utilized to provide an alternative conceptual system to make sense of the world—a system that insisted that nothing existed apart from basic particles of matter and that all phenomena could be explained in terms of the organization and the interaction of these particles.

One of the most fervent of late 19th-century materialists, T. H. Huxley, described human beings as "conscious automata" with no free will. As he explained in 1874, "Volitions do not enter into the chain of causation…. The feeling that we call volition is not the cause of a voluntary act, but the symbol of that state of the brain which is the immediate cause."

This was a very early formulation of an idea that has become commonplace among modern scientists and philosophers who hold similar materialist views: that free will is an illusion. According to Daniel Wegner, for instance, "The experience of willing an act arises from interpreting one's thought as the cause of the act." In other words, our sense of making choices or decisions is just an awareness of what the brain has already decided for us. When we become aware of the brain's

actions, we think about them and falsely conclude that our intentions have caused them. You could compare it to a king who believes he is making all his own decisions but is constantly being manipulated by his advisers and officials, who whisper in his ear and plant ideas in his head.

Many people believe that evidence for a lack of free will was found when, in the 1980s, scientist Benjamin Libet conducted experiments that seemed to show that the brain "registers" the decision to make movements before a person consciously decides to move. In Libet's experiments, participants were asked to perform a simple task such as pressing a button or flexing their wrist. Sitting in front of a timer, they were asked to note the moment at which they were consciously aware of the decision to move, while EEG electrodes attached to their head monitored their brain activity.

Libet showed consistently that there was unconscious brain activity associated with the action—a change in EEG signals that Libet called "readiness potential"—for an average of half a second before the participants were aware of the decision to move. This experiment appears to offer evidence of Wegner's view that decisions are first made by the brain, and there is a delay before we become conscious of them—at which point we attribute our own conscious intention to the act.

Iif we look more closely, however, Libet's experiment is full of problematic issues. For example, it relies on the participants' own recording of when they feel the intention to move. One issue here is that there may be a delay between the impulse to act and their recording of it—after all, this means shifting their attention from their own intention to the clock. In addition, it is debatable whether people are able to accurately record the moment of their decision to move. Our subjective awareness of decisions is very unreliable. If you try the experiment yourself—and you can do it right now, just by holding out your own arm and deciding at some point to flex your wrist—you'll become aware that it's difficult to pinpoint the moment at which you make the decision.

"If one imagines that, for me to decide something, I have to have willed it with the conscious part of my mind. Perhaps my unconscious is every bit as much 'me.' "

—Iain McGilchrist

An even more serious issue with the experiment is that it is by no means clear that the electrical activity of the readiness potential is related to the decision to move and to the actual movement. Some researchers have suggested that the readiness potential could just relate to the act of paying attention to the wrist or a button rather the decision to move. Others have suggested that it only reflects the expectation of some kind of movement, rather being related to a specific moment. In a modified version of Libet's experiment (in which participants were asked to press one of two buttons in response to images on a computer screen), participants showed readiness potential even before the images came up on the screen, suggesting that it was not related to deciding which button to press.

Still others have suggested that the area of the brain where the readiness potential occurs—the supplementary motor area, or SMA—is usually associated with imagining movements rather than actually performing them. The experience of willing is usually associated with other areas of the brain (the parietal areas). And finally, in another modified version of Libet's experiment, participants showed readiness potential even when they made a decision not to move, which again casts doubt on the assumption that the readiness potential is actually registering the brain's "decision" to move.

A further, more subtle, issue has been suggested by psychiatrist and philosopher Iain McGilchrist. Libet's experiment seems to assume that the act of volition consists of clear-cut decisions, made by a conscious, rational mind. But McGilchrist points out that decisions are often made in a more fuzzy, ambiguous way. They can be made on a partly intuitive, impulsive level, without clear conscious awareness. But this doesn't necessarily mean that you haven't made the decision.

As McGilchrist puts it, Libet's apparent findings

are only problematic "if one imagines that, for me to decide something, I have to have willed it with the conscious part of my mind. Perhaps my unconscious is every bit as much 'me.' " Why shouldn't your will be associated with deeper, less conscious areas of your mind (which are still you)? You might sense this if, while trying Libet's experiment, you find your wrist just seeming to move of its own accord. You feel that you have somehow made the decision, even if not wholly consciously.

Because of issues such as these—and others that I don't have space to mention—it seems strange that such a flawed experiment has become so influential, and has been (mis)used so frequently as evidence against the idea of free will. You might ask: Why are so many intellectuals so intent on proving that they have no free will? (As philosopher Alfred North Whitehead pointed out ironically, "Scientists animated by the purpose of proving themselves purposeless constitute an interesting subject for study.")

This is probably because the nonexistence of free will seems a logical extension of some of the primary assumptions of the materialist paradigm—such as the idea that our sense of self is an illusion and that consciousness and mental activity are reducible to neurological activity. But as I suggest in my book *Spiritual Science,* it is entirely possible that these assumptions are false. The mind may be more than just a shadow of the brain, and free will may not be an illusion but an invaluable human attribute, which can be cultivated and whose development makes our lives more meaningful and purposeful.

SCIENTIFIC AMERICAN
MIND

LETTERS TO THE EDITOR:
Scientific American, 1 New York Plaza, Suite 4600, New York, NY 10004-1562, 212-451-8200 or editors@sciam.com.
Letters may be edited for length and clarity. We regret that we cannot answer each one.

HOW TO CONTACT US:
For Advertising Inquiries: Scientific American, 1 New York Plaza, Suite 4600, New York, NY 10004-1562, 212-451-8893, fax: 212-754-1138
For Subscription Inquiries: U.S. and Canada: 888-262-5144, Outside North America: Scientific American, PO Box 5715, Harlan IA 51593, 515-248-7684, www.ScientificAmerican.com

For Permission to Copy or Reuse Material From Scientific American: Permissions Department, Scientific American, 1 New York Plaza, Suite 4600, New York, NY 10004-1562, 212-451-8546, www.ScientificAmerican.com/permissions. Please allow three to six weeks for processing.