

WHAT ARE THE LIMITS OF SCIENTIFIC EXPLANATION?

Sara Gottlieb and Tania Lombrozo

Introduction

Mary is a brilliant scientist who specializes in human vision. Her mind is so able, and her knowledge so comprehensive, that she knows all the physical facts there are to know about the perception of color. She knows exactly how light of different wavelengths travels through the environment, how it affects our retinæ, and what happens in our brains when we see different colors. She could describe the firing of every neuron and how it relates to what people report that they see. And yet, poor Mary herself has never seen the color blue (or red, or yellow ...). Due to unspecified forces, she has experienced the world from the confines of a black and white room, with her only access to the outside world provided through a black and white monitor.

One glorious day, Mary emerges from her black and white chamber to the outside world. For the first time, she sees a blue sky, a red flower, a yellow bird. As the world's expert on color vision, she already knows precisely how each surface affects the wavelengths that bounce from it, and how her brain responds to the corresponding stimulation on her retina. Yet she has never had the first-hand experience of observing these colorful entities with her very own eyes. In having this experience, does Mary learn something new about color? Or, as an expert with knowledge of all the scientific facts about color, did she already know all there is to know?

This famous thought experiment by the philosopher Frank Jackson (1986) motivates a compelling intuition: that some things can only be known through personal experience. While Mary knows everything there is to know about the science of color, there is something she doesn't know. Before emerging from her chamber, she doesn't know *what it is like* to see blue (or red, or yellow ...).

Jackson's thought experiment is usually framed in terms of the "physical information" that Mary does and doesn't know, but the example also suggests that there may be some types of knowledge that fall beyond the scope of science. Perhaps this seems obvious – science cannot, after all, tell us what values we ought to have, or what sorts of behaviors are morally good. Science is a *descriptive* enterprise, not a *prescriptive* one. But Jackson's thought experiment is powerful because it suggests that even something like color vision – a descriptive matter that vision scientists like Mary are able to study empirically – might fall beyond the scope of science. For if the first-person perspective that comes from actually *experiencing* color teaches Mary something new – *what it is like* to see blue, or red, or yellow – then there must be some kinds of knowledge about human color vision that cannot be derived from physical (or scientific) facts alone.

Those who accept Jackson's argument (and not everyone does) face a difficult choice. One possibility is to radically change the way we think about scientific knowledge to broaden its scope. More precisely, our notions of science and scientific knowledge could be expanded to include the kind of first-person knowledge that comes from first-hand experience: the *what-it's-like* to see blue. But it's not really clear how this would work. For Mary, it would mean a rejection of the premise that she can know all "scientific facts" about color from the confines of her black and white chamber. When she emerges from her chamber, she would gain new "scientific" knowledge. This challenges the way we normally think about science as an enterprise concerned with objective and verifiable knowledge – the sort of knowledge that can be captured in textbooks or formal models.

A second possibility is to accept that scientific knowledge is limited in an important respect. On this view, a complete scientific explanation for human color perception leaves something out: it doesn't supply Mary with what she needs to know *what it's like* to experience color. Correspondingly, first-person experience can supply something that falls beyond the scope of science – something that cannot be captured by a scientific explanation.

The Perceived Limits of Science

Philosophers and scientists have debated the correct response to Jackson's argument, with no clear consensus. But anecdotal evidence suggests that for many people, the view that first-person experience can supply something that falls beyond the scope of science is especially compelling. Some phenomena – especially some *psychological* phenomena – just seem to elude a scientific explanation. Could science ever really explain romantic love, altruistic behavior, or religious faith? *Should* science even be concerned with these seemingly ineffable aspects of the human condition? Perhaps scientific explanations are, in principle, limited in these domains, and perhaps this limitation is a good thing.

One source of evidence that (some) people find a view along these lines compelling comes from the experience of researchers studying romantic love (Hatfield

2006). In the mid-1970s, psychologists Elaine Hatfield, Mary Utne O'Brien, and Jane Traupmann Pillemer were awarded a small grant by the National Science Foundation for their research on passionate love and sexual desire. They were also awarded a "Golden Fleece Award" by U.S. Senator William Proxmire, who claimed that they were "fleecing" taxpayers with their research. A press release explained: "not even the National Science Foundation... can argue that falling in love is a science" (Hatfield 2006). He also opposed the research because he didn't *want* the answer: "I believe that 200 million other Americans want to leave some things in life a mystery" (Hatfield 2006). Proxmire urged the NSF to leave love to Elizabeth Barrett Browning and Irving Berlin.

Other examples of perceived limits to science come from the domains of religion and spiritual experience. In a 2012 commentary published in the journal *Nature*, for example, author Daniel Sarewitz compared the discovery of the Higgs boson to the experience of visiting the Angkor temples in Cambodia. After describing the powerful sense of mystery and transcendence elicited by the temples, alongside the sense of a universe that evades comprehension, he wrote that: "Science is supposed to challenge this type of quasi-mystical subjective experience, to provide an antidote to it." Religion can offer "an authentic personal encounter with the unknown," whereas the Higgs is "an incomprehensible abstraction, a partial solution to an extraordinarily rarified and perhaps always-incomplete intellectual puzzle." Sarewitz concludes by suggesting that "whereas the Higgs discovery gives me no access to insight about the mystery of existence, a walk through the magnificent temples of Angkor offers a glimpse of the unknowable and the inexplicable beyond the world of our experience" (Sarewitz 2012, p. 431). The upshot is that there are some things science cannot, and perhaps should not, aim to provide – a personal encounter with the unknown, or insight into the mystery of existence, chief among them.

From Anecdote to Science

Considering these examples of people's reactions to science shifts us from the realm of philosophy to the realm of human psychology. Why might (some) people have the intuition that (some) aspects of human experience, such as color perception, romantic love, or transcendent awe, fall beyond the scope of scientific knowledge? What governs which phenomena are seen as falling beyond this scope, and which within? And do these views have implications for people's attitudes towards science or scientific explanations?

The perspectives voiced by Proxmire and Sarewitz could reflect a deep truth about what scientific knowledge can and cannot do. If they are right, then scientific knowledge has real limits that both producers and consumers of science should acknowledge. Perhaps there are important aspects of color perception or romantic love or religious experience that science cannot and should not explain. Understanding such limits would be important in directing the enterprise of

science, and in recognizing the complementary contributions of other human endeavors, be they poetry or religion.

On the other hand, it could be that these intuitions about the scope of science are deeply misguided. Perhaps science *can* explain all facets of human experience, and perhaps there are important benefits that would arise from its success. On this view, perspectives like Proxmire's could impede the production or uptake of scientific knowledge. If people oppose research on sexual desire and romantic love, for example, that could interfere with the development of interventions to improve relationships or resolve sexual dysfunction.

To evaluate the basis for these intuitions, it's important to move beyond anecdote to evidence. Instead of considering potentially unrepresentative perspectives voiced in popular media, we can turn to science itself. Specifically, what does psychological science tell us about the nature and sources of people's intuitions about the scope of scientific knowledge and the limits of scientific explanation?

We decided to find out. In a series of empirical studies (Gottlieb & Lombrozo 2018), we investigated whether people in fact find some phenomena – such as love or spirituality – less amenable to scientific explanation, and we evaluated several hypotheses about why this might be the case. In the rest of this chapter, we explain what we found. But first, we consider some hypotheses that motivated our approach.

Motivating Hypotheses: Intuitive Dualism and Our Creaturely Selves

Why might people be inclined to regard some aspects of human experience as falling beyond the scope of science? One hypothesis is that people are "intuitive dualists," on some level committed to the Cartesian idea that we have minds or "souls" that are wholly different from material bodies. This view is called *intuitive* dualism because the claim isn't that people have explicitly worked-out ideas about the way the mind and the body relate, the way Descartes did, but rather that on a more intuitive or gut level, they act as if minds and bodies are fundamentally different sorts of things – the former accessible through introspection and reasoning; the latter extended in space such that we can measure and prod.

With a view like this, it makes sense that science can only offer adequate explanations for our material bodies – for our headaches, but not for our heartaches. Developmental psychologist Paul Bloom, a prominent advocate for this hypothesis about human cognition, has argued that dualist tendencies are often at odds with what science has to tell us about the physical and mechanistic substrates of the mind (Bloom 2004). At the same time, these tendencies can also help explain the allure of neuroscientific explanations for the mind. "We intuitively think of ourselves as non-physical," writes Bloom, "and so it is a shock, and endlessly interesting, to see our brains at work in the act of thinking" (Bloom 2006).

One piece of evidence for people's "intuitive dualism" comes from a clever study by Preston, Ritter, and Hepler (2013). They had people read about the psychology of love, but only some people were given additional information about the *neuroscience* of love. They read, for example, that "the ventral tegmental area and the medial caudate nucleus, associated with other forms of reward and motivation, are activated when thinking about a romantic partner." After reading this information, participants in this latter group, compared to those in the former group, reported decreased belief in a human soul or spirit. This suggests that the notion of a human soul is, at least to some extent, believed to be at odds with a reductive, scientific understanding of the human mind. In explaining the *brain*, we can't also be explaining the soul. If anything, we are explaining it away.

If intuitive dualism underlies people's resistance to the idea that science can explain human experience, then we might expect such resistance to be particularly apparent for aspects of the mind or behavior typically associated with a soul. For example, we might be more reluctant to accept a scientific explanation for why people act altruistically or fall in love than for face recognition or forgetfulness.

A related hypothesis is that for some people, the idea of being "merely" creatures – nothing more than a part of the biological world – threatens a sense of what it means to be human. According to a psychological proposal known as "Terror Management Theory," humans' awareness of their own mortality can provoke great anxiety. As a means of assuaging this anxiety, individuals tend to respond to reminders of their own mortality by employing tactics that inhibit or ease these mortality-related thoughts (Pyszczynski, Solomon, & Greenberg 2003). For example, death-related thoughts have been shown to increase the extent to which people report having religious beliefs (including belief in an afterlife, which offers a kind of immortality), even among people who do not identify as strongly religious.

Based on these ideas, we hypothesized that some scientific explanations could be "terror"-inducing. Specifically, research has shown that emphasizing our animal nature – or our own "creatureliness," as it is called within this body of research – can be perceived as threatening because it reminds us of our own mortality, triggering the terror management of Terror Management Theory (Goldenberg, Pyszczynski, Greenberg, Solomon, Kluck, & Cornwell 2001). In particular, scientific explanations that account for human traits in physical and reductionist terms, or in a way that holds equally for other species, could be rejected in an effort to manage the existential threat that they induce. This rejection could be especially robust for aspects of the mind that are perceived to make humans special. For instance, people might be less inclined to accept scientific explanations for religiosity or language, which are often perceived to be uniquely human, than for aspects of depth perception or motor control, which we share with other species to a greater extent. We refer to this hypothesis as *human exceptionalism*.

Initial Evidence that Intuitive Dualism Guides Intuitions about the Scope of Science

The two hypotheses we've introduced – that people are (to some extent) intuitive dualists, and that people are (to some extent) threatened by humans' "creatureliness" – are empirical hypotheses about the human mind. Accordingly, we can use the methods of psychological science to test them, and that's precisely what we did (you can read about these findings in more detail in Gottlieb and Lombrozo 2018).

In our first study, we presented over 300 participants with a variety of psychological phenomena, including perceiving color, experiencing "love at first sight," and having a spiritually transformative experience. But we also included phenomena that we typically associate with lower-level perceptual or cognitive processes, such as remembering somebody's name, recognizing another person's face, and reaching for and grabbing an object. There were 48 phenomena in total, and for each one, we asked: to what extent do you agree that science will one day provide a complete explanation for this phenomenon? As we expected, people were much more likely to say that science could explain phenomena that we typically associate with lower-level perceptual or cognitive processes than phenomena such as experiencing "love at first sight," or having a spiritually transformative experience.

Our next goal was to understand *why*. If people are, as psychologists have theorized, intuitive dualists, and the relevant demarcation is between bodies and minds or souls, then we would expect phenomena that involve minds or souls to be the ones most often considered beyond the scope of science. But which phenomena are these? We expected them to be those that people seem to direct or assess with their minds – that is, those over which we think we have conscious control, and those that involve an experiential quality that we can access through introspection.

To test these ideas, we had the participants in our studies again look at the same list of 48 phenomena that we presented them initially, but this time we asked them to make two novel judgments. The first was about the extent to which they considered each to involve conscious will, or the ability to deliberately influence how, when, or why the phenomenon happens. Participants tended to rate phenomena like decision-making high in conscious will, but phenomena like dreaming low in conscious will. For the second judgment, we asked a question reminiscent of Jackson's point about what Mary learned when she escaped from her chamber: "To what extent does this involve a subjective experience (a feeling of what it is like) that only the individual experiencing it can know?" For this question, participants tended to give phenomena like falling in love and believing in God high ratings, as well as acting altruistically, feeling love toward one's children, and having a sense of personal identity that persists over time. Phenomena like perceiving depth and identifying sounds received much lower ratings.

Overall, participants' judgments were consistent with the predictions of intuitive dualism: they were more likely to say that a phenomenon was beyond the scope of science if they rated that experience as requiring conscious control and as having a personal, experiential component. Phenomena that were rated high on both dimensions – such as experiencing love towards one's children or acting altruistically – were thus among the most likely to be deemed resistant to a complete scientific explanation.

These initial results echo the intuition that is so often elicited by Jackson's thought experiment about Mary. There are some mental processes or experiences – such as seeing color – that have a first-personal, experiential quality associated with them, such that a purely scientific description seems to fall short. Our results also resonate with some of the public responses to science that we quoted above. Elaine Hatfield and her colleagues received pushback when they used scientific tools to study romantic love, which rates particularly highly in having a personal, experiential component. Daniel Sarewitz argued that the Angkor temples offer “an authentic personal encounter with the unknown” (Sarewitz 2012, p. 431) – the kind of spiritual experience that our participants similarly viewed as personal and experiential, and as more likely to fall beyond the scope of science compared with many of the other phenomena that we tested.

Digging Deeper: What's So Special about Personal, Subjective Experience?

As a next step, we decided to dig deeper into what it is about having an experiential, personal, and introspectively-accessible experience that leads people to judge a phenomenon beyond the scope of science. Recall that we asked participants to answer the following question: “To what extent does this involve a subjective experience (a feeling of what it is like) that only the individual experiencing it can know?” For a philosopher of mind, this single question packs in several potentially distinct components. First, there's the matter of a *subjective experience* – the “what it is like” to experience some phenomenon. Second, there's the idea that the experience is somehow personal and *privileged* – that only the individual experiencing it can know. And finally, there's the implication that the form of access to this experience is through *introspection*, an examination of one's own thoughts and feelings. Which of these components was driving participants to judge some phenomena beyond the scope of science? Or was it all three?

In a follow-up study, we teased apart these three components of our original question. We again had participants tell us the degree to which they thought that science could ever fully explain each of the mental phenomena from our initial studies, but this time they rated those same experiences on three additional dimensions: subjective experience (“This has a subjective experience associated with it: a ‘feeling’ of what it is like”), privileged access (“Only an individual him or herself can know that he or she is experiencing this; an outside observer might be

able to guess but can't truly know”), and introspection (“An individual having this experience can know he or she experiences it through introspection: the examination of one's own internal feelings or reflection”).

In this follow-up study, we found that the latter two components – privileged access and introspection – were driving the intuition that some phenomena cannot be captured by a scientific explanation. In other words, the phenomena that participants rated as highly privileged (“only I could know!”) and as accessible through introspection were the ones rated least likely to be fully explained by science. To illustrate with an example, this suggests that people are dissatisfied with scientific explanations because falling in love has an experiential quality to it that is accessible through introspection and *only* to the experiencer herself.

It is important to note that the three dimensions we were interested in here – subjective experience, privileged access, and introspection – are all highly related to one another. However, subjective experience – the dimension related to phenomenology, or the degree to which people feel that an experience has a distinctive feeling of what it “is like” – did not have a statistically significant association with scientific explanation judgments when we statistically controlled for the other two dimensions. This could be surprising in light of Jackson's example about Mary. What she seems to learn upon first seeing color is precisely the phenomenological component of color perception, the “what it is like.” On the other hand, it makes sense that this experience is inaccessible to science, and that this is so precisely because of the way in which we access phenomenology ourselves: through the private process of introspection. An individual can introspect about her own experience; a scientist cannot do the introspecting for her, and Mary cannot “introspect” her way to the experience of seeing color by reading scientific papers or conducting research on the color perception of others.

A Role for Human Exceptionalism, Too

The evidence presented in the preceding two sections provides some support for our hypothesis about intuitive dualism: people are more resistant to the idea that science can explain a psychological phenomenon when that phenomenon is something that we take ourselves to control with our minds – through conscious will – or access with our minds – through private introspection. These are aspects of our *mental* experience – our minds, rather than our bodies.

What about our other hypothesis, linking scientific explanation to “creatureliness” and human exceptionalism? We were especially curious to test the idea that people are more resistant to scientific explanations for traits that are believed to be uniquely human, relative to those we share with other species. Recall that this hypothesis was motivated by the idea that reductionist or cross-species explanations for uniquely human traits could be threatening because they liken us to our animal relatives and remind us of our own mortality. To test this, we asked people to rate the very same list of 48 mental phenomena from the studies

already described, but this time we had them indicate whether they thought they were uniquely human, or present in other species as well. Things like falling in love, making moral judgments, and religion were considered uniquely human, but so were a variety of complex cognitive tasks, such as engaging one's imagination and thinking creatively. Other phenomena, such as dreaming and integrating sensory information to figure out where a sound is coming from, received low ratings for human uniqueness.

Consistent with our prediction, we found that the phenomena rated high on human uniqueness were also more likely to be judged beyond the scope of science. That is, a phenomenon related to hearing or seeing – examples of perceptual processes – was more likely to be judged amenable to a “complete scientific explanation” than a phenomenon like making moral judgments.

In a follow-up study, we dug deeper into the idea of human uniqueness by unpacking two separate components that could have contributed to people's judgments. We ultimately found that people were particularly resistant to the idea that science could explain things that contribute to making humans *exceptional* relative to other species. So it isn't just that imagination and creativity are present *only* in human minds (or that people believe them to be so), but that the ability to exhibit these characteristics is perceived to be part of what makes humans special. If this is correct, then fully explaining imagination or creativity (versus motor control or depth perception) in scientific terms could seem implausible because it's taken to imply a fully physical or reductionist account of the capacity. It folds us into the biological realm – a mere creature among many – and fails to set humans apart from other species. And if Terror Management Theory is right, being a mere creature is an uncomfortable reminder of our own mortality.

Our studies thus provide some support for two initial hypotheses – that people are (to some extent) intuitive dualists, and that scientific explanations are (on average) judged to be less likely for our less-creaturely, more uniquely human characteristics. But these are just two of many possible hypotheses. We also tested a third hypothesis: that people might treat *complexity* as a fundamental constraint on scientific knowledge. That is, people might consider something like romantic love to be beyond the scope of scientific explanation because it is perceived to be too complex, and in particular more complex than basic cognitive or perceptual processes. The most interesting thing about this hypothesis is that we robustly failed to find any support for it. We found that romantic love *was* considered highly complex, but so were things like logical reasoning and memory. More crucially, these complexity ratings were *unrelated* to judgments about the possibility of obtaining a complete scientific explanation for the corresponding phenomenon.

What Should Science Explain?

Summarizing the findings we've just described, our studies revealed that some phenomena are typically judged to resist a complete scientific explanation, and

that science is not perceived to be limited by the complexity of its subject matter, but instead by its third-personal and potentially reductive methodology. These findings speak to public skepticism about the idea that science could one day fully explain romantic love or transcendent awe. But they don't yet speak to another aspect of our introductory examples: the sense that there would be something *bad* about achieving a complete scientific explanation; that when it comes to some things, science *should* be limited. Recall Proxmire's admonition that “Americans *want* to leave some things in life a mystery” (emphasis added), and Sarewitz's implication that science shouldn't purport to offer more than it does – to quote the title of his piece, “sometimes science must give way to religion.” Do most people share this sense that some scientific explanations are not only impossible, but also unwelcome?

To find out, each of our studies also asked people to tell us how *uncomfortable* they would be if science could fully explain the phenomenon in question. Interestingly, people were most uncomfortable with the idea that science could explain things like love, morality, or religious belief – the very same things they said that science could never *possibly* explain. Also mirroring our initial results, we found that people were uncomfortable with science explaining things they felt they could consciously will, and things that made humans exceptional compared to other species. Moreover, these judgments related to ratings of privileged access and introspection in exactly the same way as the scientific possibility questions did: people were most uncomfortable with the idea that science could fully explain the phenomena that they deemed knowable only by the experimenter herself, and those that supported introspective access.

Why do scientific explanations for some phenomena generate discomfort? It could be that intuitive dualism and creatureliness are at work once again, but this time manifesting in a more visceral form, making us *uncomfortable* with the very idea that science could succeed when it comes to explaining our uniquely human minds. But these findings also raise interesting questions about the source of people's beliefs concerning what science *should* or *should not* seek to explain – the sorts of beliefs that affect the research scientists choose to pursue, the projects that funding agencies choose to support, and the public's response to their efforts. Beyond practical considerations, are these beliefs about what science should and shouldn't pursue largely governed by the suite of epistemic and affective responses that our studies reveal? If so, our findings have important implications, as they uncover subtle aspects of human psychology that shape the course of science.

Some Open Questions

So far we've been talking about averages – how people respond, on average, to questions about whether science could possibly provide a full explanation for a given phenomenon, and about whether such an explanation would be

uncomfortable. Our studies also raise some questions about differences across individuals that warrant further investigation.

In our original studies, we surveyed a rather diverse online sample of individuals who had a range of educational backgrounds. Surprisingly, our findings replicated fully in a sample of undergraduate students who had taken, on average, a handful of psychology courses, suggesting that commitments about the appropriate scope of science are fairly stable despite modest scientific training. It remains unclear, however, whether professionally-trained cognitive scientists would demonstrate the same pattern of results.

That said, there did exist some variation among individuals in the degree to which they thought science could or should explain the mind. And although these differences did not differ systematically with education, they did correlate with political ideology and religiosity: people who thought that science cannot and should not explain aspects of the human mind were more likely to be politically conservative and religious. Was it conservatism and religiosity that led to intuitive dualism and exceptionalism, or the other way around? This is an important question for another day.

Another open question concerns the possibility that despite resistance to the *idea* of a complete scientific explanation, such an explanation – once offered – might actually be accepted, and even welcome. Recall psychologist Paul Bloom's observation that despite finding brain-based explanations for the mind unintuitive, we find them "endlessly interesting" (Bloom 2006).

In fact, there's evidence that people *like* explanations for psychological phenomena that appeal to neuroscience. Specifically, research has found that people are susceptible to what is called the "reductive allure" effect: they prefer explanations at lower levels (e.g., that appeal to neuroscience) to explanations at higher levels (e.g., that appeal only to psychology), even when the lower-level content does not offer additional explanatory information (Hopkins, Weisberg, & Taylor 2016). Take, for example, the psychological phenomenon known as "the other-race effect," which shows that people have difficulty telling two faces apart when those faces come from a race other than their own. In the "reductive allure" studies, participants were asked to evaluate a psychological explanation for the effect. For half of those participants, the psychological explanation did not appeal to neuroscience:

In communities where the majority of the people are white, white faces are seen more frequently than are those of other races. This greater experience with white faces tunes the perceptual system to recognize greater detail across those faces, making it easier to tell them apart.

For the other half, the psychological explanation was more reductive in that it included additional neuroscientific information: "This greater experience with white faces tunes the fusiform face area to recognize greater detail across those faces, making it easier to tell them apart." Participants without relevant expertise

generally thought that the latter explanation was a better one, even though expert participants did not.

So could it be that our participants were simply wrong about their anticipated discomfort, and that scientific explanations for all mental phenomena will in fact be welcome as soon as they're on offer? We suspect not. When it comes to evaluating explanations for the types of phenomena that tend to fall beyond the scope of science – such as romantic love or religious experience – the allure of reduction could be offset by the allure of intuitive dualism, and by the repulsion of our "creaturely" selves. Our dualism and exceptionalism hypotheses would both predict that, despite the reductive allure, people will be more uncomfortable with explanations of love, for example, as those explanations become increasingly reductive. That is, people would be more uncomfortable with a chemical explanation than a neuroscientific one, and more uncomfortable with a neuroscientific explanation than a psychological one. This is a question for future research.

Implications for the *Real* Limits of Scientific Explanation

It's important to emphasize that the body of research we've been discussing reveals *intuitions* about what science can or cannot explain, and that it does not speak directly to what science can, in fact, explain. How, then, should these results be interpreted? On the one hand, it could be that people's intuitions track some epistemic truth about the limits of science. If these intuitions are correct, then despite methodological advances, science will never fully explain something like romantic love because of its rich, first-personal and uniquely human experiential quality. On the other hand, if these intuitions are instead misguided, they could prove to be serious barriers to scientific advance, leading people to have intuitive biases against scientific explanations in certain domains. Should this be the case, there is the concern that scientists could avoid or fail to receive support for research in areas that many consider outside the scope of science, even when that research could lead to important theoretical and practical advances.

Going one step further, if people falsely believe that a scientific perspective is not only *insufficient*, but also misplaced or even harmful, we could miss out on potentially important truths about ourselves and the world. For example, many believe that gene editing technologies, such as CRISPR-Cas9, hold the promise of transforming medicine by eliminating previously incurable diseases and disorders. In 2015, the journal *Science* referred to it as the "breakthrough of the year." But many surrounding discussions, including those by scientists working on such technologies, have focused on whether we *should* be implementing these technologies in the first place, and for what purposes. George Daley, a stem cell researcher and the dean of Harvard Medical School, remarked on a 2017 success in using gene-editing technology to alter viable human embryos, saying, "The question now remains should we – and for what purposes and should there be certain applications that are allowed and others that are prohibited?" (Marrin 2017):

As society moves forward in both debating and embracing advances in gene editing, it will be important to query public opinion: which types of applications do people generally consider acceptable, and which types of applications do people generally consider unacceptable? Are people more opposed to intervening on some traits than others? These are open empirical questions, but it might be that people are more uneasy with gene editing when it plays upon some of our commitments about the limits of science – for example, when it aims to target traits perceived as uniquely human, or ones typically associated with a human soul or essential spirit. These are important questions to be addressing – and it's important to get the answers right, even when they might violate initial intuitions.

Thus returning from the realm of human psychology to the realm of philosophy, we can ask with more urgency: When it comes to the scope of science, are people's intuitions getting things right, or getting things wrong? We think the answer is "both."

Regarding intuitive commitments about introspection and privileged first-person access, we think people's intuitions might be onto something important. Perhaps there really are, in principle, certain aspects of experience that cannot be captured by scientific knowledge alone. This is one of the points illustrated by the case of our black-and-white Mary who has the scientific knowledge, but not the experiential knowledge, of seeing color. Science benefits from its objective, third-person methodology, and this methodology will one day allow us to explain why Mary does or doesn't have the experiences that she does. But these explanations will supply scientific knowledge, not personal experience.

Regarding intuitive commitments about human exceptionalism and our desire to be more than mere creatures, we're more inclined to dismiss intuition. If resistance to scientific explanations for uniquely human traits is motivated by mortality-related anxiety and existential threat, it's not clear whether or why these judgments might also track some epistemic truth. However, it could well be that scientific explanations fall short of providing everything we want. It's not that they fail to fully explain a phenomenon, but rather that they don't put it in a personal and cultural context that reflects its human significance. For that we may well benefit from the arts and humanities, from poetry and music. We agree with Daniel Sarewitz that "[t]he Higgs boson, and its role in providing a rational explanation for the Universe, is only part of the story." (Sarewitz 2012, p. 431).

Coda

We can now imagine the subject of a new thought experiment, Mary-Lou. Mary-Lou is a college student taking her first philosophy course. Her professor has just assigned Jackson's piece about Mary the color scientist and asks the students to ponder whether there are limits to scientific knowledge. That night, she sits in her dorm room reading about Mary stepping out of her black and white chamber for the first time. Consistent with the data we've presented, Mary-Lou thinks back to

her professor's question and intuitions that scientific knowledge is limited in its scope. Does the case of Mary-Lou the philosophy student demonstrate that there are true limits to scientific knowledge? On its own, the answer is surely no: intuitions are not always correct. But intuitions are often a first step; armed with data and arguments, they can sometimes show us the way.

References

- Bloom, P., 2004. *Descartes' baby: How the science of child development explains what makes us human*, New York: Basic Books.
- Bloom, P., 2006. Seduced by the flickering lights of the brain. *Seed Magazine*, 27.
- Goldenberg, J.L., Pyszczynski, T., Greenberg, J., Solomon, S., Kluck, B. and Cornwell, R., 2001. I am not an animal: Mortality salience, disgust, and the denial of human creatureliness. *Journal of Experimental Psychology: General*, 130(3), p. 427.
- Gottlieb, S. and Lombrozo, T., 2018. Can science explain the human mind? Intuitive judgments about the limits of science. *Psychological Science*, 29(1), pp. 121–130
- Hatfield, E., 2006. The Golden Fleece Award: Love's labours almost lost. *APS Observer*, 19(6): 16–17.
- Hopkins, E.J., Weisberg, D.S. and Taylor, J.C.V. 2016, The seductive allure is a reductive allure: People prefer scientific explanations that contain logically irrelevant reductive information. *Cognition*, 15, 67–76.
- Jackson, F., 1986. What Mary didn't know. *Journal of Philosophy*, 83(5), pp. 291–295.
- Maron, D. 2017. Embryo gene-editing experiment reignites ethical debate. *Scientific American*, [online] Available at: www.scientificamerican.com/article/embryo-gene-editing-experiment-reignites-ethical-debate/ [Accessed 7 April 2018].
- Preston, J.L., Ritter, R.S. and Hepler, J., 2013. Neuroscience and the soul: Competing explanations for the human experience. *Cognition*, 127(1), pp. 31–37.
- Pyszczynski, T., Solomon, S. and Greenberg, J., 2003. *In the wake of 9/11: Rising above the terror*. Washington, DC: American Psychological Association.
- Sarewitz, D., 2012. Sometimes science must give way to religion. *Nature News*, 488(7412), p. 431.