

UC Berkeley

UC Berkeley Electronic Theses and Dissertations

Title

Can we explain the human mind and experience? Perceptions of the limits of science

Permalink

<https://escholarship.org/uc/item/5vp3w838>

Author

Gottlieb, Sara

Publication Date

2018

Peer reviewed|Thesis/dissertation

Can we explain the human mind and experience?
Perceptions of the limits of science

By

Sara G Gottlieb

A dissertation submitted in partial satisfaction of the
requirements for the degree of

Doctor of Philosophy

in

Psychology

in the

Graduate Division

of the

University of California, Berkeley

Committee in charge

Professor Tania Lombrozo, Chair
Professor Alison Gopnik
Professor Dacher Keltner
Professor Wesley Holliday

Summer 2018

© 2018 – Sara Gottlieb
All rights reserved

Abstract

Can we explain the human mind and experience?
Perceptions of the limits of science

by

Sara Gottlieb

Doctor of Philosophy in Psychology

University of California, Berkeley

Professor Tania Lombrozo, Chair

Can science ever fully explain what it is like to fall in love, to undergo a religious transformation, or to have a child? For some types of experiences, people seem to give special status to their own introspective knowledge over other types of information, especially scientific information, an inherently third-personal source.

Across these three sets of studies, I demonstrate that people privilege their own first-person knowledge over science when it comes to highly personal aspects of the human mind and experience. In Chapter 2, I show that people judge science as being unable to ever explain aspects of the mind to which they feel that they have privileged introspective access – the types of mental phenomena that only the experiencer herself can truly know. In Chapter 3, I extend this finding to show that these commitments hold important consequences for people’s moral judgments. Finally, in Chapter 4 I demonstrate that people privilege their own first-person knowledge over scientific information when making decisions about personally transformative decisions.

These findings advance the literature on folk epistemology, but also illustrate that epistemic commitments have important consequences for moral judgments and how people approach decision-making in their everyday lives.

Acknowledgements

This dissertation – and the other work I completed during graduate school – would not have been possible without the support and mentorship of my advisor, Tania Lombrozo. She has helped me grow as a scientist and as an individual, and for that I feel grateful beyond words. Tania is a brilliant researcher with an incredibly sharp mind, and my career will be forever shaped by her insight and guidance. I have come to appreciate that not all advisors invest in their students to the extent that Tania does; her ability to support us – both professionally and personally – is truly admirable.

I am also deeply appreciative of my committee members – Alison Gopnik, Dacher Keltner, and Wesley Holliday – who provided helpful feedback on my research and helped spark new and interesting ideas, and also Laurie Paul who provided very helpful feedback on my transformative experience work. I also thank my lab mates who have provided useful comments over the years: Caren Walker, M. Pacer, Dylan Murray, Daniel Wilkenfeld, Nadya Vasilyeva, Carly Giffin, Emily Liquin, Dillon Plunkett, and Ellie Kon. I also appreciate the great friends and work buddies that spent countless hours writing at cafes (and sharing both laughs and tears) with me: Antonia Peacocke, Jackson Kernion, Rachel Jansen, Mariel Goddu, and Jess Jones. These people kept me sane and happy, and I miss you all already.

Many thanks also go to John Schindel, who was a pillar of support during my time at Berkeley. John is somewhat of a wizard with his wealth of knowledge about all things Berkeley – and is unwaveringly kind, understanding, patient and caring despite his many, many responsibilities. There are many times that I (and every single other graduate student) would have been lost without his help. I also thank the rest of the support staff at Berkeley, and also the National Science Foundation, the McDonnell Foundation, and the John Templeton Foundation for providing financial support for this research.

I could not have gotten to this point without the support of my entire family – my parents, my three older siblings, and my nieces and nephews. Being across the country from you all was not easy; California sure is nice, but I'm happy to be back in New England and in close proximity to (most of) you.

The biggest thanks, however, goes to Andrew, who is not yet my husband but will be by the time that this is published. He is my biggest cheerleader, and my partner in all things.

Finally, I'd like to thank my dog, Moose, for cuddling by my side as I wrote roughly 80% of this dissertation.

Table of Contents

Chapter 1. Introduction	1
Intuitive commitments about the limits of science	2
Do some types of knowledge fall beyond the scope of science?.....	3
Can science serve as a good guide for decision-making?.....	5
Goals of the present dissertation.....	5
Chapter 2. Can science explain the human mind? Intuitive judgments about the limits of science.....	9
Study 1	9
Study 2a: Introspection/phenomenology	14
Study 2b: Human uniqueness	14
Study 2c: Abnormal functioning	15
Study 3a	16
Study 3b.....	17
Discussion.....	18
Chapter 3. Dissociable components of mind-body dualism predict bioethical attitudes	20
Study 1	22
Study 2	25
Study 3	27
Study 4	29
Study 5	30
Study 6	34
Discussion.....	36
Chapter 4. Choosing when we cannot know: First-person priority in transformative choice	38
Study 1a	39
Study 1b	41
Study 2	43
Study 3	45
Discussion.....	47
Chapter 5. Discussion	49
Do some types of knowledge fall beyond the scope of science?.....	49
Can science serve as a good guide to decision-making?	52
Implications for the <i>real</i> limits of science	54
References.....	56

List of figures

Figure 2.1 Mean scientific possibility and comfort ratings for six representative items.....	11
Figure 2.2 Predictors used in Studies 1-3a.....	13
Figure 3.1 Results from Study 5	34
Figure 3.2 Results from Study 6	36
Figure 4.1 Average “know” judgments from Study 1a.....	41
Figure 4.2 Average “decide” judgments from Study 1b.....	42
Figure 4.3 Results from Study 3	47

List of tables

Table 2.1 Zero-order correlations in Study 1.....	12
Table 2.2 Results of the two regression models conducted in Study 3a.....	17
Table 2.3 Zero-order correlations in Studies 3a and 3b.....	18
Table 2.4 Results of the two regression models conducted in Study 3b.....	18
Table 3.1 Summary of the Dualism Complex scale and its five components.	23
Table 3.2 Factor loadings of the 32-item Dualism Complex scale.....	25

Chapter 1. Introduction

Mary is a brilliant scientist who specializes in human vision. Her knowledge is so comprehensive that her colleagues regard her as knowing 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...). For some unspecified reason, 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.

On one brilliantly sunny day, Mary is finally afforded the freedom to emerge from her black and white chamber. She steps outside for the first time and looks up to see the blue sky. She looks around her to see a red flower and a yellow bird. As she revels in the beauty of her colorful new world, has Mary learned something new about color? Or, as an expert with knowledge of all the physical facts about color and human vision, did she already know all there is to know?

This famous thought experiment by the philosopher Frank Jackson (1986) motivates a compelling intuition: some things can only be known through personal experience. Despite the fact that Mary is said to know everything there is to know about the science of color, there is something she doesn't know. Before emerging from her black and white room, she does not know *what it is like* to see red, or blue, or yellow; she critically lacks the first-personal type of information that seems so essential for truly knowing color.

This thought experiment is usually framed in terms of the physical information that Mary does and does not know. It also suggests, however, that there may be some types of knowledge that fall beyond the realm of scientific explanation, or, alternatively, that some types of knowledge can only be truly known through firsthand experience itself. It may seem trivial that science genuinely cannot answer a wide range of questions. Science cannot, for example, tell us which values we ought to hold, or which actions and behaviors are most worthy of praise or punishment. Science prides itself on being a descriptive enterprise, not a prescriptive one. But this thought experiment gains so much traction because it elicits an even more surprising intuition. It should not be taken as so trivial that something like color vision – a mental phenomenon that scientists like Mary can and do study empirically, using the scientific method – might also fall beyond the scope of science. Science is concerned with physical facts about the world; these are facts that can be learned and verified from a third-person perspective, and these are the facts that Mary has learned through her extensive scientific training. Therefore, if Mary's wondrous experience of finally seeing color does, in fact, teach her something new – which our intuitions quite strongly tell us it does – then there must exist some non-physical facts about human color vision, or, at the very least, some kinds of knowledge about human color vision, that cannot be derived from the physical facts alone.

This thought experiment also raises a second set of normative and empirical questions regarding decision-making. When Mary is granted the freedom to emerge from her chamber, she must decide whether or not to leave the familiar safety of her black and white room. For many decisions involving our future selves, we imagine ourselves undergoing new experiences, and what they would be like *for us*. But for something so epistemically transformative as seeing color

for the first time – and with access only to scientific knowledge, as Mary does – it seems that her ability to imagine her future self in a colorful world might be limited or even impossible. How should Mary, therefore, make a rationally informed decision about whether to stay or leave? This scenario may appear a bit contrived – especially since Mary would be eager to finally experience color, the subject to which she has devoted her entire career – but a more relatable quandary does, in fact, arise frequently in our daily lives. How should an individual rationally decide, for example, whether or not to take a mind-altering drug, or to have a child for the first time, when she is unable to anticipate what those transformative experiences will be like? Can (and should) science guide people to making appropriate decisions in these cases, or does it leave something out?

These two points – one about what science can or cannot explain, and one about the role science plays in making decisions involving our future selves – both center around the question of whether science can fully capture the “what it’s like” aspect of certain mental states and human experiences. Many aspects of the mind and how we, as individuals, perceive and see the world are inherently private experiences. An individual herself knows what it is like to fall in love or feel awed by nature in a way that is not apparent to an outside observer. She might tell other people what it is like, and other people can study her outward behavior, but the richness of her subjective inner experience is available only *to her*.

This dissertation presents a series of studies investigating people’s folk epistemic commitments broadly related to this issue, and more specifically investigates the degree to which people perceive science as limited in domains within which they tend to privilege their own first-person knowledge. Chapters 2 and 3 are motivated by the first question raised above: do some types of knowledge fall beyond the scope of science? More precisely, do people judge science as being unable to explain certain aspects of the mind but not others? Chapter 4 is motivated by the second question: when do people think that science can serve as a good guide to decision-making? Do people privilege their own first-person knowledge when making decisions involving their future selves?

Chapters 2 and 3 focus specifically on people’s perceptions of the limits of science when it comes to explaining the human mind. In Chapter 2, I ask why it is that people generally consider topics like romantic love and religious belief, but not topics like memory or motor control, as beyond the scope of scientific explanation. Chapter 3 investigates the relationship between these epistemic commitments and moral judgments, and also situates these commitments within the literature on intuitive dualism. Finally, Chapter 4 addresses the second question by asking people to judge the extent to which science, as well as other sources of information, should help them in deciding whether or not to engage in new experiences.

Intuitive commitments about the limits of science

It is important to recognize that this dissertation presents a series of experiments investigating people’s intuitions about the limits of science and scientific explanation. It does not reveal the *actual* limits of science, and should not be taken as support for or against arguments like the one laid out by Jackson’s (1986) thought experiment. It should come as no surprise that not all philosophers support the use of experimental methods to investigate folk intuitions. Some have argued, for example, that philosophers do not and should not rely on intuitions as evidence for their views (Cappelen, 2012). And others have argued that folk notions of concepts – knowledge, for example – do not actually uncover the nature of that concept in the real world (Kauppenin, 2007). Nonetheless, experimental work on folk epistemology can have important

consequences for how we think about existing debates within philosophy. For example, past work on people's perceptions of what constitutes knowledge versus belief has important theoretical implications for debates within epistemology (Myers-Schulz & Schwitzgebel, 2013).

Uncovering and explaining folk epistemic intuitions has important consequences for how people regard science and scientific explanations in their everyday lives. If people do see science as being seriously limited in its scope, then understanding such limits would be important in directing the enterprise of science and science communications, and also in recognizing how other fields – be they social sciences, like sociology or anthropology, or even art and literature – can contribute valuable and complementary types of knowledge.

Further, answering the two questions that motivate this project has important consequences for judgment and decision-making. In Chapter 3, I show that people's epistemic commitments are related to their stances on bioethical issues such as abortion and cloning, which are divisive across a wide political spectrum and among the most controversial issues in contemporary United States politics. I show that people who view science as limited in its ability to explain the human mind tend to be more opposed to not only these issues, but also stem-cell research and physician-assisted suicide, and that this relationship holds independently of religiosity and political orientation. Understanding the ways in which epistemic commitments and moral judgments are intertwined can help shed light on the origin and development of a certain set of sociopolitical beliefs, and also raises important issues about how scientific advances may influence societal level views on bioethical debates.

Gaining a deeper knowledge of these epistemic commitments also aids our understanding of how people weigh science, and other sources of information, when making everyday decisions. In Chapter 4, I show that people generally privilege their own first-person knowledge over science when making decisions about their future selves, especially when these decisions are highly personal and transformative in nature. This finding provides some insight into why scientific evidence can be devalued in people's decision-making procedures, and can even inform future efforts to promote evidence-based decision making in high stakes cases, be they medicine or issues of public policy.

In what follows, I discuss in more detail the two motivating questions outlined above, as well as more thoroughly lay out the goals of the present dissertation.

Do some types of knowledge fall beyond the scope of science?

Do people perceive science as limited in its ability to explain the human mind? Many of our everyday experiences are inherently private, available to us through the process of introspection. Despite the fact, for example, that Mary the color scientist can study color from an “outside” perspective, it is tempting to say that she does not know what it is like to perceive color “inside” of her head. And outside the world of fictional thought experiments, people commonly judge that not all mental phenomena can be fully explained or understood through science, an outside, third-person perspective (e.g., Hatfield, 2006; Sarewitz, 2012).

Anecdotally, there has been public resistance to scientific studies on a range of topics related to human experience – including romantic love, religious belief, and moral judgment – but it is unclear why. People often think that science cannot, or should not, explain these highly personal and sacred topics, but the same is not true of more basic cognitive and perceptual properties, like motor control or the ability to recognize faces. In the case of romantic love, for example, a reverend criticized researchers in the 1970's, asking, “Who granted these ‘scientists’ the ability to see into men's minds and hearts?” (Hatfield, 2006).

To date, little research in psychology has directly assessed people's folk epistemic commitments about the limits of science. However, when it comes to how people think about the role of science in explaining the mind, some inspiration comes from work on intuitive mind-body dualism – the theory that humans naturally carve the world into material bodies and immaterial souls. Descartes famously held that the mind and body, or the mind and brain, are two radically different kinds of things (Robinson, 2017), with the human mind or spirit being free from the physical laws that govern our material bodies. Psychological or intuitive dualism is most often associated with the developmental psychologist Paul Bloom, who argues that people, from a very early age, navigate the world as young Cartesians (Bloom, 2009). He does not suggest that adults (or children) have explicitly worked-out metaphysical theories about the way the mind and the body or the mind and brain relate to one another, as philosophers do. Rather, he argues that people are intuitively committed to the idea that the mind and the brain are not the same thing; while the former is accessible through introspection, the latter is extended through space and accessible to third-person individuals, including scientists.

Bloom argues that one domain in which these intuitive commitments manifest is in people's views on science and scientific explanations. Specifically, he and others (e.g., Greene, 2011; Farah & Murphy, 2009) suggest that our dualist tendencies are often at odds with what science has to tell us about the physical and mechanistic substrates of the human mind. Farah and Murphy (2009), for example, explain that a neuroscientific understanding of phenomena such as love, morality, and spirituality – but not things like visual perception or memory – undermines the intuitive notion of a ghost in the machine (Ryle, 1949). At the same time, however, these tendencies might also explain why neuroscience, as a field, is so alluring. Bloom (2006) explains, “We intuitively think of ourselves as non-physical, and so it is a shock, and endlessly interesting, to see our brains at work in the act of thinking.”

Despite that fact that dualist commitments are purported to be universal (Slingerland & Chudek, 2011; Chudek, McNamara, Birch, Bloom, & Henrich, 2013), it is nonetheless theorized that they vary in strength among individuals. As a result, there do exist several scale measures aimed at tracking individual differences in dualist commitments. These scales tend to include items related to a suite of themes commonly associated with Cartesian dualism – including, but not restricted to, views on mind-brain identity, religious commitments to a soul, belief in an afterlife, beliefs in free will and determinism, and, most relevant to this discussion, science and reductionism. For example, the Stanovich (1989) Dualism Scale includes statements such as, “Knowledge of the mind will forever be beyond the understanding of sciences like physics, neurophysiology, and psychology.” A more recent scale (Preston, Ritter & Hepler, 2013) includes similar types of items (e.g., “Aspects of the mind that science cannot explain are best explained by the soul”). These researchers also found empirical evidence for the claim that people experience an inherent tension between their intuitively dualist tendencies and neuroscientific explanations; they found that people who read about neuroscientific explanations for phenomena like romantic love (as opposed to psychological ones) went on to report a decreased belief in a human soul or spirit.

If intuitive dualism does underlie people's resistance to the idea that science can explain human experience, then we would expect this tension to be particularly apparent for aspects of the mind that are typically associated with the soul – aspects that involve conscious will, can be accessed through introspection, or are accompanied by a distinctive phenomenology or subjective experience. For example, this literature predicts that people would be more reluctant to accept a scientific explanation for why people act morally praiseworthy, which typically

involves conscious will, than for how or why people dream, which typically does not. This hypothesis is tested directly in the studies reported in Chapter 2, and in Chapter 3 I explore how these epistemic commitments relate to moral judgments, as well as intuitive dualism more broadly.

Can science serve as a good guide for decision-making?

The case of Mary the color scientist suggests to us that not only are some topics considered beyond the scope of scientific knowledge, but also that some experiences – like seeing red, or blue, or yellow – can only truly be known through firsthand experience, and pose a tricky puzzle for accounts of rational decision-making. Mary might know about color through her extensive scientific training, and she might know how other people describe their own experiences of seeing color, but to say that she really *knows* color, in some deeper or richer sense, can seem unsatisfying. Despite her extensive scientific training, she is unable to anticipate the phenomenological, subjective experience of seeing red, or blue, or yellow (Nagel, 1974).

The experience of seeing color for the first time is considered to belong to a special set of cases referred to as transformative experiences. The philosopher L.A. Paul (2014) suggests that having a child, undergoing sex reassignment surgery, or (in an even more extreme case) becoming a vampire are all potentially transformative in the sense that they have the capacity to change a person's epistemic position and identity: the individual gains access to new experiential information as a consequence of the choice itself, and in some sense becomes a different person.

Epistemologists find transformative experiences particularly interesting because, among other reasons, it is near impossible for individuals to anticipate the subjective experience of what those experiences would be like for their future selves. Further, these experiences, as Paul argues, also pose a special puzzle for rational decision-making (Briggs, 2017). The pre-choice individual, the argument goes, cannot anticipate the subjective value of each potential outcome, and is subsequently unable to rationally choose between options. In becoming a vampire, for example, how should one choose whether or not to undergo the transformation if she cannot actually know what being a vampire will be like?

This is, to some extent, a challenge posed by all decisions; no experience is completely identical to one of the past, and so simulating what any future event will be like is an imperfect procedure. Further, people are notoriously bad at predicting their affective responses to future events (Wilson & Gilbert, 2013). Nonetheless, the puzzle of transformative choice highlights an important epistemic dilemma and raises a series of empirical and normative questions about how people think they can know what experiences will be like, and critically how they think they ought to decide whether or not to engage in them. People may judge science as limited in its ability to inform their decision-making procedures when subjective, first-personally assigned value is considered especially relevant. This issue is approached empirically in Chapter 4.

Goals of the present dissertation

The present dissertation consists of three sets of studies, each investigating distinct aspects of people's folk epistemic commitments concerning the limits of science and scientific explanation. Across these studies, I will show that people tend to privilege their own first-person knowledge over science, an inherently third-person source of information, when it comes to certain aspects of the human mind and experience. In Chapters 2 and 3 – the chapters on scope of science commitments – this will manifest in people perceiving science as limited in its capacity

to fully explain aspects of the mind to which they feel that they have privileged, first-person access. In Chapter 4 – the chapter on decision-making – this will manifest in people devaluing science, and instead privileging their own first-person knowledge, when making choices involving their future-selves, especially those decisions that are considered personally transformative.

Chapter 2 assesses people's intuitions about the scope of science – namely, whether people think that fully scientific explanations insufficiently capture the nature of some psychological phenomena but not others. In a series of studies, I present individuals with a list of psychological phenomena and ask them to indicate the degree to which science could ever fully explain each, and also ask them to rate each on dimensions I hypothesize as relevant to scope of science judgments. In Studies 1-3, I find that individuals are more likely to judge scientific explanations as impossible when those phenomena support first-person, introspective access (e.g., feeling empathetic as opposed to reaching for objects), contribute to making humans exceptional (e.g., appreciating music as opposed to forgetfulness), and involve conscious will (e.g., acting immorally as opposed to having headaches). Interestingly, I find that these epistemic judgments are also closely linked to non-epistemic judgments: people generally report discomfort with scientific explanations for these same psychological phenomena. These results did not vary as a function of education; in Studies 3a and 3b, I report similar results from both an online sample and an undergraduate sample with formal training in psychology.

One important finding from this set of studies is a null result. In Study 1, I ask individuals to rate the perceived complexity of each of these phenomena, and I do not find evidence that these complexity ratings predict judgments concerning the appropriate scope of science. This surprising result suggests that the perceived limits of science are not determined by the perceived complexity of its subject matter, but instead by the third-personal, non-intentional perspective that is characteristic of scientific methodology. These results speak directly to the first question posed at the outset of this dissertation: people do indeed perceive some types of knowledge as beyond the scope of science. When it comes to the human mind, people consider science as limited in its ability to explain the types of experiences or mental states where first-person knowledge is considered especially relevant, such as those perceived as accessible only through introspection.

In Chapter 3 I explore the link between these scope of science commitments and intuitive dualism more broadly, and also investigate the relationship between these commitments and moral judgments. In particular, I look at judgments concerning sanctity of life bioethical issues – including abortion, embryonic stem-cell research, and cloning. Dualism is often invoked in conversations about why people oppose these types of bioethical issues; abortion, for example, can be thought of as a debate over when a human embryo becomes host to a human soul. This link has been discussed in past research both theoretically (Bloom, 2004; Greene, 2011) and empirically (Richert & Harris, 2008). The goals of Chapter 3 are to 1) better understand the construct of intuitive dualism by decomposing it into its constituent parts, and 2) use these dissociable elements to predict bioethical judgments.

In Study 1, I create a novel dualism scale called the Dualism Complex. Interestingly, I find that commitments concerning the relationship between the mind and the brain – the aspect most strongly associated with metaphysical dualism – only account for a relatively small amount of variance in judgments on the Dualism Complex scale. Instead, religious commitments to a soul, and also judgments concerning the appropriate scope of science in explaining the mind, emerge as considerably larger factors. In Study 2 I use these components to predict judgments on

a range of bioethical issues. When controlling for belief in god and political conservatism, two individual differences strongly related to both dualist attitudes and bioethical judgments, I find that only scope of science commitments predict bioethical attitudes. This interesting result suggests that it is a very narrow aspect of dualism – namely, scope of science judgments – that relates to judgments regarding sanctity of life issues.

In the next several studies, I aim to better characterize this relationship and its causal structure. I fail to find evidence for a straightforward causal relationship between scope of science commitments and bioethical attitudes. Instead, the final studies reveal that these two may share a set of common determinants. More specifically, I find that that when the mind is perceived as having certain qualities – such as the ability to exert conscious will or the ability to introspect on its own mental activity – people tend to perceive it as being beyond the scope of scientific explanation (as I also find in Chapter 2), and they also tend to judge bioethical issues, such as cloning or stem-cell research, as more morally impermissible when involving such entities. Taken together, these findings clarify that there does indeed exist a relationship between dualism and bioethical attitudes – and one that exists independent of religiosity or political orientation – though it is specifically scope of science commitments, an aspect of dualism that deviates from metaphysical dualism in the stricter sense.

Chapter 4 explores people’s folk epistemic commitments about decision-making, including 1) how one can come to know what an experience will be like, and 2) how one ought to decide whether or not to engage in that experience. On an everyday basis, we make an endless number of decisions – some rather trivial, and some less so – but we often are unable to anticipate what those experiences will be like, just as Mary the color scientist is unable to anticipate the experience of seeing color despite her extensive scientific training. It might seem trivial that people will rely on a certain source of information for their decision-making – testimony, for example – to the extent that it supports knowledge about what that experience will be like, but recent philosophical work on transformative experience suggests that this is likely not the case. Indeed, the series of studies I present suggest that people hold normative commitments strongly privileging first-personal sources of information when making decisions, despite the fact that they perceive first-personal sources of information as limited in their ability to tell them what new experiences will be like.

In Study 1a, I ask people to make judgments about how they can know what a range of scenarios will be like. Half of these are canonical transformative experiences (e.g., having a child), and half are not (e.g., having a knee replacement), and I ask people to what extent they think that science, testimony, and their own imaginations can tell them what each experience will be like. In Study 1b, people make similar judgments, but this time indicate to what degree they think they ought to rely on science, testimony, and their own imaginations when making the hypothetical decision whether or not to engage in each experience. I find an interesting mismatch between these two sets of judgments: people generally rate testimony as the best source of information for knowing what an experience will be like, but judge imagination as most important when making decisions. This pattern is true for both transformative and non-transformative decisions, but most pronounced for the former. For these cases, people judge imagination as especially compromised for knowing, but especially important for deciding.

In Study 2, I ask whether it is personal transformations, epistemic transformations (or both) that drive the observed effect. To investigate this, I ask people to rate a range of experiences on the degree to which each would transform them personally and epistemically, and also ask them to answer the same “knowing” and “deciding” questions from Studies 1a and 1b.

Additionally, in order to account for the fact that transformative decisions are often more consequential decisions than non-transformative ones, I control for the stakes, or relative importance, of each experience and decision. I find that epistemic transformations drive judgments about how people can come to know what experiences are like, but that personal transformations drive people's unwillingness to privilege non first-personal sources of information when deciding whether or not to engage in that experience. Further, these results remain even while controlling for stakes.

In the final study, I ask people to report how they made one transformative or non-transformative choice in their own lives. I largely replicate findings from the previous studies, suggesting that normative commitments about decision-making mirror actual decision-making strategies in the real world.

Interestingly, personally transformative experiences often possess many of the same characteristics that I find, in Chapter 2, to influence scope of science judgments. More specifically, transformative experiences – which are often considered highly individual, sacred experiences – have a distinctive phenomenology or subjective experience associated them, are perceived as being accessible only to the experiencer herself, and contribute to making humans exceptional compared to other species. Common examples of personally transformative experiences include falling in love, religious experiences, or feeling awed by a beautiful landscape, and these are among the same types of experiences that I find people to judge as being beyond the scope of scientific explanation.

Taken together, the three sets of studies presented in this dissertation focus on people's perceptions of the limits of science, and, in addition to providing theoretical advance in folk epistemology, show that these commitments have important consequences for moral judgment and everyday decision-making. People often judge their own mental states and inner experiences as available only *to them*, which is reflected in a commonly held perception of science being unable to fully capture what it is like to experience a range of mental phenomena, often ones that are deeply personal, transformative, and phenomenologically rich. The studies reported in this dissertation are among the first to empirically investigate these intuitions.

Chapter 2. Can science explain the human mind? Intuitive judgments about the limits of science

Research on romantic love has had a complicated past. In 1975, senator William Proxmire awarded Elaine Hatfield and her colleagues a “Golden Fleece Award” for their research on love, claiming they were “fleecing” taxpayers with “unneeded” research. A reverend echoed what many were thinking: “Who granted these ‘scientists’ the ability to see into men’s minds and hearts?” (Hatfield, 2006). These reactions reflect intuitive beliefs about the scope of science – beliefs about which questions *cannot* or *should not* be approached scientifically. But where do these beliefs come from? Why are falling projectiles appropriate targets of scientific research, while falling in love is not?

To our knowledge, these questions have not been investigated empirically, but prior work hints at a class of phenomena that could be commonly regarded as falling beyond the scope of science: phenomena associated with the mind or the soul. This proposal has roots in philosophy (e.g., Robinson, 2016), but psychologists have subsequently argued that humans are natural-born dualists, carving the world into minds and bodies. Bloom (2004), for instance, has suggested that dualist tendencies are often at odds with what science has to tell us about the physical substrates of the mind (see also Preston, Ritter, & Hepler, 2013), which could account for resistance toward scientific research on topics like love.

The present research had two goals. First, we sought to chart people’s beliefs about the appropriate scope of science when it comes to explaining the human mind. To do so, we asked participants about the *possibility* of scientific explanations for various psychological phenomena, ranging from romantic love to more basic perceptual processes. Participants also indicated whether they were *uncomfortable* with the idea that science could fully explain each phenomenon. These judgments provided insight into participants’ epistemic commitments as well as the non-epistemic judgments intimated by Proxmire and others – that there is something unsettling about scientific explanations for certain phenomena, whether this judgment stems from an affective response or personal values.

The second goal of our research was to identify what it is that differentiates psychological phenomena perceived to fall within versus beyond the scope of science. Inspired both by philosophical discussions of mind-body dualism and empirical work on free will (Bloom, 2004; Shariff et al., 2014; Nichols & Knobe, 2007; Nadelhoffer, Shepard, Nahmias, Sripada, & Ross, 2014), we expected that phenomena with an introspectively accessible phenomenology, and over which we have some conscious will, would be more likely to fall beyond the perceived scope of science. Based on research in moral psychology and personal identity, we also anticipated that a phenomenon would be more likely to fall beyond the perceived scope of science if regarded as unique to humans (Haslam, Bain, Douge, Lee, & Bastian, 2005; Goldenberg, Pyszczynski, Greenberg, Solomon, Kluck, & Cornwell, 2001) and central to identity (Strohinger & Nichols, 2014; 2015). Finally, we predicted that phenomena that suggest abnormal functioning would be more likely to fall within the scope of science (Plunkett, Lombrozo, & Buchak, 2014), and we measured the perceived complexity of each phenomenon as a variable that could plausibly affect the perceived possibility of a complete scientific explanation.

Study 1

Method: We recruited 317 participants from Amazon Mechanical Turk in exchange for payment (155 females; mean age = 35 years, $SD = 11$ years; 46% reported a college degree or

higher). For all studies, we filtered out individuals who had participated in conceptually related studies from our lab group. This also prevented an individual from participating in more than one study reported in this paper. All participants had Mechanical Turk approval ratings greater than 95%, which has been shown to ensure high data quality (Peer, Vosgerau, & Acquisti, 2014). As such, no participants in any of our studies were excluded from analyses, though data from participants who did not complete the study were not analyzed.

Participants rated 46 mental traits, abilities, or phenomena (presented in a random order) on two ratings of scientific explanation: “Science could one day fully explain ___” (scientific possibility) and, “I am uncomfortable with the idea that science could one day fully explain ___” (scientific discomfort). All scientific possibility ratings were made on one page, and all scientific discomfort ratings were made on a second page, but the order of these pages was randomized. Scientific discomfort scores were reverse coded as “scientific comfort” in order to trend in the same direction as scientific possibility. Sample items included falling in love, reaching for objects, and using language to communicate (see Figure 1.1 for additional items and their corresponding scientific possibility and comfort ratings; a complete list of items can be found on OSF at <https://osf.io/3r96f/>). Items were selected to span a range of psychological topics, from perception and language to morality and emotion. We identified six dimensions on which we expected the phenomena described by these items to vary. Participants were randomly assigned to rate each item on *one* of these six dimensions (italicized titles of dimensions not presented to participants):

Introspection/phenomenology: “___ involves a subjective experience (a feeling of what it is like) that only the individual experiencing it can know.”

Human uniqueness: “___ is unique to humans.”

Abnormal functioning: “___ indicates abnormal functioning.”

Conscious will: “People have conscious will over ___ - they can deliberately influence when, how, or why it happens.”

Important for identity: “___ is important for identity; it is a central aspect of what makes up a person’s true self.”

Complexity: “___ is complex.”

Participants indicated their agreement with all statements on a scale from 1 (“strongly disagree”) to 7 (“strongly agree”). These ratings always followed the ratings for scientific possibility and discomfort, with items once again presented in a random order on a single page.

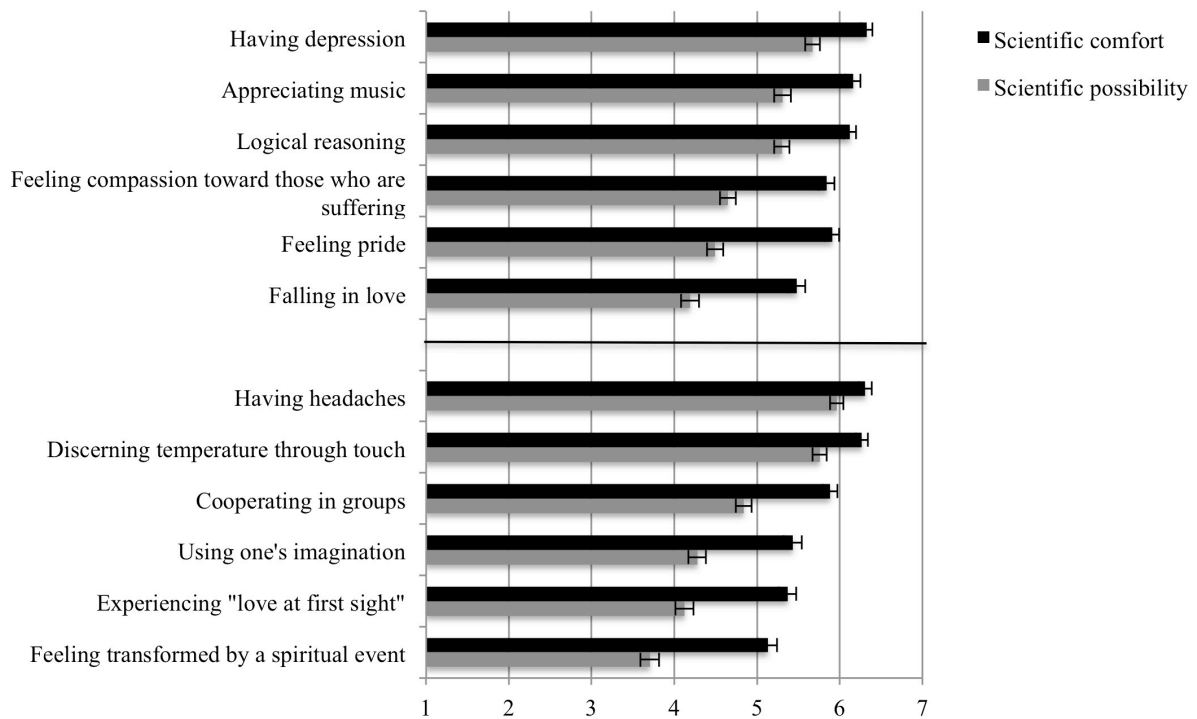


Figure 2.1 Mean scientific possibility and comfort ratings for six representative items (from a total of 46) from Study 1 (top), and for six representative items (from a total of 92) from Study 3 (bottom). The items presented on top were used in Studies 1-3, while those presented on the bottom were used in Study 3 only. Error bars represent one standard error above and below the mean.

Results: Our sample size ensured that roughly 50 participants provided ratings for each dimension (as is the case for all studies that follow). We chose this number to exceed the sample size used by Bear and Knobe (2016), who conducted similar analyses and assigned 30 participants to each of their conditions. We also exceeded their number of stimulus items (30).

Because rating dimension was a between-subjects factor, we analyzed relationships between our dimensions and dependent variables across items ($n = 46$). We first created means for each item's scientific possibility and scientific comfort ratings using data from all participants. We then created means for each item corresponding to the six rated dimensions, in each case using data from the subset of participants who rated the corresponding dimension. Scientific possibility and scientific comfort were each correlated with all dimensions, with the exception of complexity (see Table 1.1 for full correlation matrix). Scientific possibility and scientific comfort were also highly correlated with each other ($r = .92$). Our subsequent analyses therefore considered these ratings both in conjunction and individually.

We next sought to identify the *unique* variance contributed by each dimension. We therefore included our six dimensions as predictors in a multivariate analysis that included both scientific possibility and scientific comfort as outcome variables. We found significant multivariate effects of introspection/phenomenology, $F(2,38) = 60.02, p < .001$, human uniqueness, $F(2,38) = 6.97, p = .003$, abnormal functioning, $F(2,38) = 6.41, p < .004$, and

conscious will, $F(2,38) = 5.74, p = .007$. There were no significant multivariate effects of importance for identity, $F(2,38) = 1.91, p = .162$, nor complexity, $F(2,38) = .28, p > .250$.

We followed this analysis with univariate models, both of which accounted for a high proportion of variance (adjusted $R^2 = .83$ and $.84$, respectively, $ps < .001$). In the model with scientific possibility as the outcome variable, we again found that introspection/phenomenology ($\beta = -.77, p < .001$), human uniqueness ($\beta = -.28, p < .001$), abnormal functioning ($\beta = .37, p = .002$), and conscious will ($\beta = -.37, p = .001$) were all significant predictors, while importance for identity and complexity were not ($\beta = .18, p = .151$ and $\beta = -.08, p > .250$, respectively). For scientific comfort, we similarly found that introspection/phenomenology ($\beta = -.78, p < .001$), human uniqueness ($\beta = -.17, p = .027$), abnormal functioning ($\beta = .37, p = .002$), and conscious will ($\beta = -.23, p = .037$) were significant predictors, while importance for identity and complexity were not ($\beta = -.01$ and $\beta = -.06, ps > .250$, respectively). In both models introspection/phenomenology was the strongest predictor of scientific explanation ratings.

For subsequent studies, we focused on the dimensions that accounted for significant unique variance in these individual models, but we additionally report correlations between each dimension and scientific possibility and comfort.

Dimension	1	2	3	4	5	6	7	8
1. Scientific possibility	1.00	0.92***	-0.67***	-0.51***	0.32*	-0.39**	-0.39**	-0.06
2. Scientific comfort		1.00	-0.74***	-0.35*	0.36*	-0.32*	-0.53***	-0.12
3. Introspection/phenomenology			1.00	0.15	0.16	-0.19	0.25	0.48***
4. Human uniqueness				1.00	0.00	0.24	-0.13	0.04
5. Abnormal functioning					1.00	-0.61***	-0.66***	0.51***
6. Conscious will						1.00	0.47***	-0.59***
7. Important for identity							1.00	-0.02
8. Complexity								1.00

Table 2.1 Zero-order correlations (across items) among predictors and dependent variables in Study 1. Note: * $p < .05$, ** $p < .01$, *** $p < .001$.

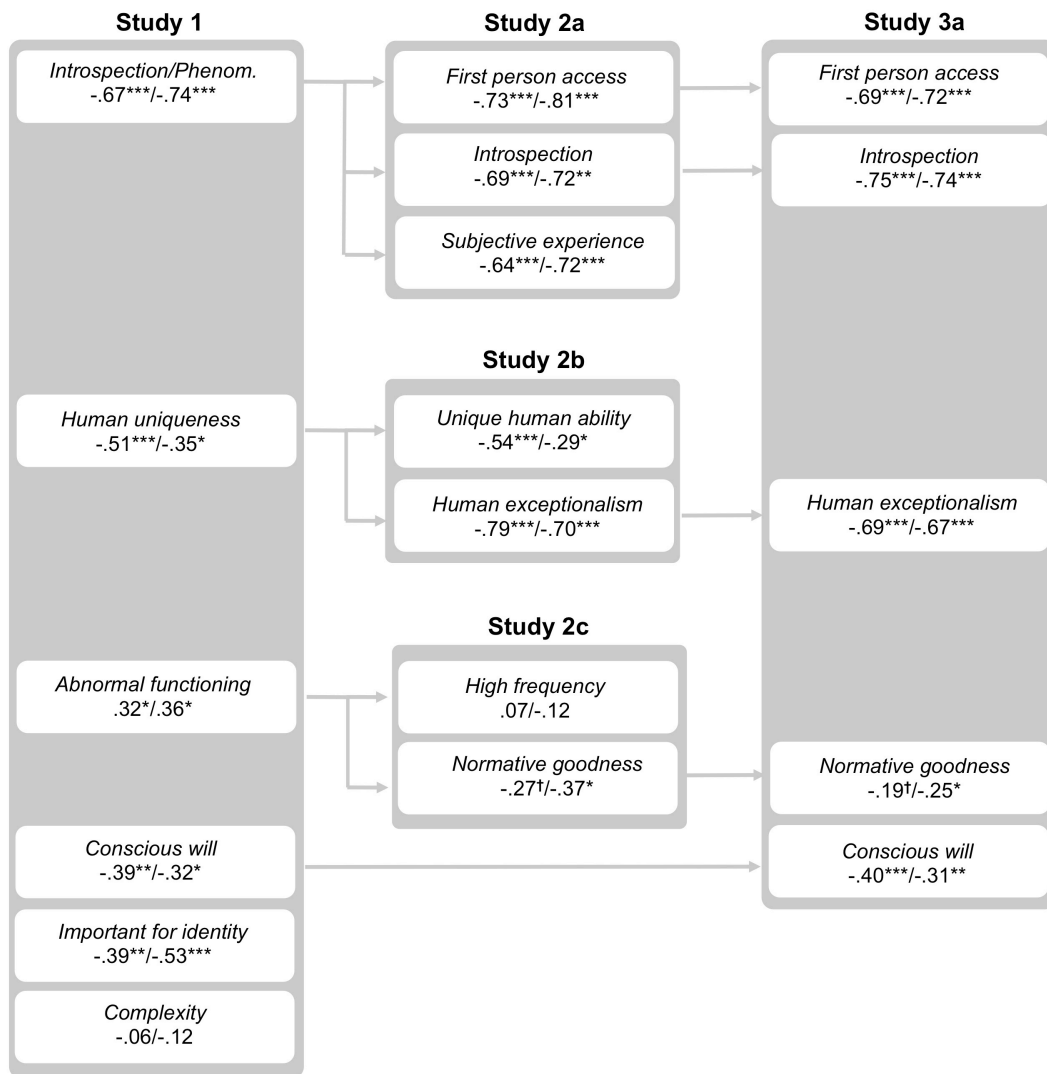


Figure 2.2 Predictors used in Studies 1-3a. The numbers below each predictor correspond to the correlation between that predictor and each of scientific possibility (first) and scientific comfort (second). Note: † $p < .07$, * $p < .05$, ** $p < .01$, *** $p < .001$.

Discussion: Study 1 isolated several dimensions relevant to the perceived scope of science in explaining the human mind: introspection/phenomenology, human uniqueness, abnormal functioning, conscious will, and importance for identity. Equally important, however, is a relationship that we did not find: participants were not simply committed to the idea that science cannot explain phenomena that are perceived to be complex. In Studies 2a-2c, we revisit three significant dimensions that involved multiple components to identify which component was responsible for the associations with scientific possibility and discomfort.

Study 2a: Introspection/phenomenology

Study 1 found that people judge scientific explanations less likely and more uncomfortable for phenomena that support first-person introspective access, or some subjective feeling. The dimension assessed in Study 1 combined several elements related to introspection/phenomenology (see Schwitzgebel, 2016); in Study 2a we tease these apart.

Method: 218 individuals were recruited from Amazon Mechanical Turk (120 females; mean age = 32 years, $SD = 11$ years; 47% of the sample had a college degree or higher).

Participants saw stimulus items identical to those used in Study 1. All participants again rated the 46 items for scientific possibility and scientific discomfort, and were then assigned to rate each phenomenon on one additional dimension that could have played a role in the introspection/phenomenology dimension identified in Study 1:

Privileged first person access: “Only an individual him or herself can know that he or she is experiencing ___; an outside observer might be able to guess but can’t truly know.”

Introspection: “An individual having the experience can know he or she experiences ___ through introspection: the examination of one’s own internal feelings or reflection.”

Subjective experience: “___ has a subjective experience associated with it – a ‘feeling’ of what it is like.”

Results: As in Study 1, we computed correlations between each of the three dimensions and the two dependent variables by using average ratings for each item (see Figure 1.2). We also used a multivariate analysis to test for the effects of our three dimensions on our two dependent variables. The multivariate analysis revealed a significant effect of privileged first person access, $F(2,41) = 7.23, p = .002$, but not introspection, $F(2,41) = 2.42, p = .102$, nor subjective experience, $F(2,41) = .51, p > .250$.

We again followed this multivariate analysis with two separate regression models, using either scientific possibility or scientific comfort as the outcome variable. In the scientific possibility model (adjusted $R^2 = .57, p < .001$), we found that privileged first person access ($\beta = -.46, p = .007$) and introspection ($\beta = -.32, p = .037$) were significant predictors, but subjective experience was not ($\beta = -.07, p > .250$). For the scientific comfort model (adjusted $R^2 = .69, p < .001$), we found similar results: privileged first person access ($\beta = -.53, p < .001$) and introspection ($\beta = -.25, p = .053$) were significant predictors, but subjective experience was not ($\beta = -.13, p > .250$).

These models suggest that subjective experience was not responsible for unique variance in scientific possibility and comfort ratings, and that privileged access may be the single most important factor. However, because all three dimensions were highly correlated with each other (all $r_s > .70$), we also conducted a dominance analysis (Budescu, 1993; Nimon, Oswald, & Roberts, 2013) to measure the relative importance of predictors in our models. For both scientific possibility and scientific comfort models, we found that privileged first person access was the dominant predictor, and that subjective experience was dominated by the other two dimensions, consistent with the results of the individual multiple regression models.

Study 2b: Human uniqueness

In Study 2b we unpacked another dimension from Study 1, human uniqueness, by differentiating between phenomena found only among humans and those that contribute to making humans exceptional.

Method: 122 participants (60 females, mean age = 33 years, $SD = 11$ years) were recruited from Amazon Mechanical Turk (57% of this sample had a college degree or higher). Participants again rated the 46 items on scientific possibility and scientific discomfort. They were then assigned to rate phenomena on one of the following two dimensions related to human uniqueness:

Unique human ability: “Only humans have the ability to do ____.”

Human exceptionalism: “The ability to ____ is part of what makes humans exceptional.”

Results: Correlations are reported in Figure 1.2. When we entered both dimensions in a multivariate regression with scientific possibility and scientific comfort as outcome variables, we found significant multivariate effects of both unique human ability, $F(2,42) = 7.76, p = .001$, and human exceptionalism, $F(2,42) = 23.54, p < .001$.

Individual models also accounted for significant variance (adjusted $R^2 = .65, p < .001$ for the model with scientific possibility, and adjusted $R^2 = .47, p < .001$, for the model with scientific comfort). In both cases we found that human exceptionalism was a significant predictor ($\beta = -.71$ and $\beta = -.74, ps < .001$, respectively), but unique human ability was not ($\beta = -.19, p = .071$ and $\beta = .07, p > .250$). Because these two predictors were highly correlated with one another ($r = .49$), we again conducted a dominance analysis to evaluate the relative importance of predictors. For both the scientific possibility and scientific comfort models, we found that human exceptionalism was the dominant predictor, consistent with the results of the individual multiple regression models.

Study 2c: Abnormal functioning

Study 2c unpacked a third dimension from Study 1 by disentangling the statistical and normative dimensions of abnormality (see, e.g., Bear & Knobe, 2017; Hitchcock & Knobe, 2009; Uttich & Lombrozo, 2010; Wachbroit, 1994). A behavior can be normal in the statistical sense that it is common, even if it is not considered to be good (e.g., jaywalking), or normal in the sense that it is considered to be good or ideal, even if it is statistically uncommon (e.g., normal weight). We untangled these dimensions by having participants rate phenomena for frequency (i.e., normality in a statistical sense) versus “goodness” (i.e., normality in a normative sense).

Method: 125 participants were recruited from Amazon Mechanical Turk (61 females; Mean age = 35 years, $SD = 12$ years; 50% reporting a college degree or higher). Participants rated the 46 items on scientific possibility and scientific discomfort. They were then assigned to rate items on one of the following two dimensions related to normality:

High frequency: “Most people are able to ____.”

Normative goodness: “It is good to be able to ____.”

Results: Correlations are reported in Figure 1.2. We entered both dimensions as predictors in a multivariate analysis with both scientific possibility and scientific comfort as outcome variables. Neither high frequency nor normative goodness had significant multivariate effects, $F(2,42) = .44, p > .250$, and $F(2,42) = 2.44, p = .100$, respectively.

Individual regression models did not account for significant variance (adjusted $R^2 = .05, p = .122$ for scientific possibility; adjusted $R^2 = .07, p = .071$ for scientific comfort), but in both cases we found that normative goodness was negatively related to acceptance of scientific

explanations ($\beta = -.39, p = .048$ and $\beta = -.42, p = .031$, respectively), while a trait's frequency was not ($\beta = .18$ and $\beta = .16$, respectively, $ps > .250$). The two predictors were highly correlated with one another ($r = .65$), so we again conducted a dominance analysis to evaluate the relative importance of predictors. For both scientific possibility and scientific comfort models, we found that normative goodness was the dominant predictor.

Study 3a

In a final study, we tested the predictive value of our significant dimensions more stringently by (a) considering additional items and (b) considering all dimensions that accounted for unique variance in Studies 1 and 2a-c in a single model.

Method: 317 individuals were recruited from Amazon Mechanical Turk in exchange for payment (151 females; mean age = 34 years, $SD = 10$; 46% reporting a college degree or higher).

We increased the number of items participants saw from 46 to 92, thereby increasing our degrees of freedom since analyses were conducted across items. Participants again rated each item on scientific possibility and scientific discomfort, but we modified the wording of these questions. In Studies 1-2, some phenomena (e.g., making moral judgments, falling in love, recognizing faces) were preceded by “why” (e.g., “Science could one day fully explain why people fall in love”). Other phenomena did not have “why” (e.g., “Science could one day fully explain the ability to use language to communicate”). To standardize the wording for all phenomena, participants rated statements with the following form:

Scientific possibility: “Science could one day fully explain the following phenomenon: _____.”

Scientific discomfort: “I am uncomfortable with the idea that science could one day fully explain the following phenomenon: _____.”

All participants completed this pair of ratings, as well as a final set of item ratings for a single dimension – one of: first person access, introspection, human exceptionalism, normative goodness, and conscious will. These dimensions were worded identically to how they appeared in previous studies, and were chosen based on the fact that they were responsible for unique variance in the individual models reported in Studies 1-2.

Results: We first calculated correlations between each of our dimensions and scientific possibility and scientific comfort. Consistent with Studies 1-2, we found significant correlations in the expected direction between all dimensions and possibility/comfort ratings ($ps < .035$), with the exception of normative goodness and scientific possibility, which was marginal ($p = .065$).

We entered our five dimensions as predictors in a multivariate analysis with both scientific possibility and scientific comfort as outcome variables. This analysis revealed significant multivariate effects of all predictors: $F(2,85) = 9.11, p < .001$ for first person access, $F(2,85) = 4.53, p = .014$ for introspection, $F(2,85) = 11.30, p < .001$ for human exceptionalism, $F(2,85) = 26.28, p < .001$ for normative goodness, and $F(2,85) = 13.58, p < .001$ for conscious will. We then tested two multiple regression models across items with all five predictors, one on scientific possibility and one on scientific comfort (see Table 2.2). All predictors remained significant and in the expected direction, with the exception of normative goodness. Surprisingly, normative goodness was *positively* associated with both scientific possibility and scientific comfort in our regression models ($\beta = .47, p < .001$ and $\beta = .28, p = .001$, respectively), although the zero-order relationships between normative goodness and scientific possibility and scientific comfort remained *negative* ($rs = -.19, p = .065$ and $-.25, p = .017$). This suggests that the partial

relationship between normative goodness and our dependent variables is actually a positive one, once the other four predictors are taken into account.

Predictor	Scientific possibility			Scientific comfort		
	<i>b</i>	<i>SE b</i>	β	<i>b</i>	<i>SE b</i>	β
First person access	-0.36	0.09	-0.35***	-0.20	0.05	-.39***
Introspection	-0.31	0.10	-0.27**	-0.16	0.06	-.28*
Human exceptionalism	-0.38	0.08	-0.43***	-0.15	0.05	-.33**
Normative goodness	0.21	0.03	0.47***	0.06	0.02	.28**
Conscious will	-0.22	0.05	-0.32***	-0.06	0.03	-.17*
R^2	.78***			.68***		

Table 2.2 Results of the two regression models conducted in Study 3a. Note: * $p < .05$, ** $p < .01$, *** $p < .001$.

Study 3b

Study 3b aimed to test the generality of the results of Study 3a by replicating the study in a different sample: undergraduate students with exposure to psychology courses.

Method: 299 individuals were recruited from the undergraduate participant pool at UC Berkeley in exchange for course credit (151 females, mean age = 21 years, $SD = 3$). On average, participants had taken four psychology courses. The procedure of Study 3b was identical to Study 3a.

Results: Correlations between each dimension and our two dependent variables are presented in Table 2.3. All pairs were significantly correlated with the exception of scientific possibility and normative goodness. A multivariate analysis identical to that in Study 3a revealed significant effects of first person access, $F(2,85) = 30.08, p < .001$, human exceptionalism, $F(2,85) = 11.42, p < .001$, normative goodness, $F(2,85) = 17.30, p < .001$, and conscious will, $F(2,85) = 10.53, p < .001$, but not introspection, $F(2,85) = 2.10, p = .128$. The individual regression models revealed significant unique effects of all predictors, with the exception of introspection on scientific comfort (see Table 2.4). However, replicating Study 3a, normative goodness was *positively* related to scientific comfort when variance due to the other four predictors was accounted for.

Finally, we asked whether psychology background influenced overall possibility and comfort ratings. We did not find a relationship between the number of psychology courses taken and overall scientific possibility ratings ($r = -.02, t(295) = -.34, p > .250$) or overall scientific comfort ratings ($r = .05, t(295) = -.84, p > .250$). Surprisingly, however, we found that our undergraduate population displayed lower overall scientific comfort ratings compared to the Mechanical Turk population used in Study 3a ($M = 5.51, SD = 1.30$ for undergraduates and $M = 5.80, SD = 1.30$ for Mechanical Turk), $t(614) = 2.79, p = .006$. There were no group differences for overall scientific possibility ratings ($M = 4.83, SD = 1.11$ for undergraduates and $M = 4.76, SD = 1.29$ for Mechanical Turk), $t(614) = 0.77, p > .250$.

Dimension	Study 3a		Study 3b	
	Scientific possibility	Scientific comfort	Scientific possibility	Scientific comfort
First person access	-0.69***	-0.72***	-0.81***	-0.84***
Introspection	-0.75***	-0.74***	-0.61***	-0.62***
Human exceptionalism	-0.69***	-0.67***	-0.72***	-0.67***
Normative goodness	-0.19 [†]	-0.25*	-0.17	-0.21*
Conscious will	-0.40***	-0.31**	-0.47***	-0.38***

Table 2.3 Zero-order correlation coefficients of five predictors and each of scientific possibility and scientific comfort in Study 3a (left) Study 3b (right). Note: [†] $p < .07$, * $p < .05$, ** $p < .01$, *** $p < .001$.

Predictor	Scientific possibility			Scientific comfort		
	<i>b</i>	<i>SE b</i>	β	<i>b</i>	<i>SE b</i>	β
First person access	-0.55	0.08	-0.53***	0.37	0.05	-0.62***
Introspection	-0.19	0.09	-0.16*	-0.08	0.06	-0.12
Human exceptionalism	-0.31	0.06	-0.34***	-0.13	0.04	-0.26**
Normative goodness	0.17	0.03	0.34***	0.05	0.02	0.20**
Conscious will	0.22	0.05	-0.26***	-0.07	0.03	-0.15*
R^2	.82***			.77***		

Table 2.4 Results of the two regression models conducted in Study 3b. Note: * $p < .05$, ** $p < .01$, *** $p < .001$.

Discussion

Some topics are perceived to be more appropriate targets for scientific research than others. To our knowledge, however, no research has examined which psychological phenomena are believed to fall beyond the scope of science, and why this is the case. Our results suggest that people are more likely to regard a psychological phenomenon as lying beyond the scope of science when that phenomenon supports privileged introspective access, makes humans exceptional, and so on. Moreover, these judgments about the epistemic scope of science are accompanied by discomfort at the idea that science could fully explain the phenomenon in question.

These judgments reveal theoretically important aspects of folk epistemology: science is not perceived to be limited by complexity itself, but by the non-intentional, third-personal perspective of scientific methodology. Judgments about what science perhaps *shouldn't* explain – as reflected in scientific discomfort – additionally speak to the relationship between epistemic commitments and values. Further investigating this relationship is an important direction for future research. It could be that people are comfortable with what science can explain, believe science can explain what they are comfortable with science explaining, or that these judgments have an overlapping set of determinants.

Importantly, the mind is not the only topic for which people resist scientific explanations; human evolution is a case in point. It seems plausible that some of our predictors – such as human exceptionalism – will extend to other scientific domains, whereas others – such as introspective access – are unique to scientific explanations for the mind. Future work is necessary to understand how judgments about the human mind sit within a larger epistemic framework, and how they relate to a broader range of scientific and bioethical issues.

Pursuing this work is important given the relevance of scientific explanations for education, public health, and beyond: scientific explanations for global warming have been shown to shift belief in anthropogenic climate change (Ranney & Clark, 2016), scientific explanations for health behaviors have been shown to generate behavior change (Weisman & Markman, 2017), and scientific explanations for criminal behaviors can shift legal judgments (e.g., Denno, 2015). We suggest that such effects could be moderated by the factors our studies reveal. For example, scientific explanations involving mental health are likely to affect judgments and behavior, but we would expect such effects to be moderated by whether the experience or behavior being explained is believed to be good, consciously willed, and so on.

Our studies also reveal that the associations we identify between scientific possibility and comfort, on the one hand, and our predictor dimensions, on the other, are not restricted to a single population. We find highly consistent results across a diverse on-line sample and a more homogenous undergraduate sample with exposure to formal education within psychology. At the same time, our findings hint at important differences across populations in the perceived scope of science; future work will seek to understand how other individual differences, such as socioeconomic status and political orientation, influence these judgments.

Finally, our findings have important implications for policy and public uptake of science, and more generally shed light on potential resistance to a fully scientific explanation of the mind and human behavior.

Chapter 3. Dissociable components of mind-body dualism predict bioethical attitudes

We experience the world as mental entities: we think, we believe, and we desire. But we also experience the world as material bodies: we have boundaries, move through space, and interact with other objects. These two ways of experiencing the world – as minds or souls on the one hand, and as bodies or brains on the other – are potentially difficult to reconcile. Indeed, some have suggested that humans are “intuitive dualists” (Bloom, 2009), who share with Descartes a deep commitment to the fundamental difference between minds and brains (Descartes, 1641).

In the current paper we investigate the relationship between intuitive dualist beliefs and sanctity of life bioethical judgments, such as those concerning the permissibility of abortion and physician-assisted suicide. Specifically, we ask whether metaphysical commitments about the relationship between the mind and brain shape people’s attitudes towards such interventions, and if so, why this is the case. Before motivating our studies, we consider prior work on dualism and its relationship to bioethics.

In the philosophy of mind, dualism refers to the idea that the mental and the physical, or the mind and body or brain, are radically different kinds of things (Robinson, 2017). Descartes, in particular, was a substance dualist. He believed that the mind, in having the capacity to think, was different in kind from physical matter. More nuanced forms of dualism exist as well. For example, a property dualist believes that, though the world is composed of only physical matter, there exist distinct physical and mental properties. Despite this variation in dualist positions, the term “dualism” is most strongly associated with a metaphysical position like Descartes’.

In the psychological literature, intuitive dualism similarly refers to the belief that the mind and the body, or the mind and the brain, are fundamentally distinct. Paul Bloom (2004), for example, argues that “we see the world in terms of material bodies, including our own bodies, and in terms of souls [...] Our dualist conception isn’t an airy intellectual thing; it is common sense and rooted in a phenomenological experience.”

Although articulations of intuitive dualism align closely with a Cartesian-style substance dualism, there has been little empirical work on intuitive dualism, and much of it is either indirect or packaged with a broader set of commitments, such as beliefs about the existence of an afterlife or free will. For instance, one task (Haidt, Bjorklund and Murphy, 2000) asks participants to “sell their soul” to the experimenter by signing a document indicating that they will sell their soul for a certain sum of money. Reluctance to sign the document is taken as a measure of belief in an immaterial mind or soul. A more recent set of studies by Forstmann, Burgmer, and Mussweiler (2012) asks participants to indicate the degree to which the mind and body are the same entity by arranging two circles – one labeled mind, and one labeled body – with some degree of overlap. Lower levels of overlap are taken to reflect a greater commitment to dualism.

Only a handful of papers have aimed to measure dualist beliefs more explicitly. One of the scales most widely used to measure dualist beliefs is from Stanovich (1989), and it includes items that go well beyond the core belief in substance dualism. For example, some items tap into beliefs about mind-brain identity (e.g., “Minds are inside brains but are not the same as brains”), but others tap into afterlife beliefs (e.g., “My consciousness will survive the disintegration of my physical body”), and yet others into science and reductionism (e.g., “When I imagine a scene in my mind, I am in a state that will forever be beyond explanation by science”). A more recent scale from Preston, Ritter, and Hepler (2013) surveys judgments about materialism and

neuroscience more narrowly, and in so doing introduces items that relate to beliefs about reductionism and the scope of science (e.g., “Aspects of the mind that science cannot explain are best explained by the soul”).

Understanding intuitive dualism is important not only because it potentially reflects our implicit metaphysical commitments, but also because dualist attitudes are assumed to affect a host of other attitudes and behaviors. In particular, dualism is often invoked in explaining public attitudes towards science, technology, and bioethics (Bloom, 2004; Greene, 2011). Greene (2011), for example, explains: “The debate over abortion is ultimately a metaphysical one. The question is not whether a fertilized egg is alive, but whether it is host to a ‘human life,’ i.e., a human soul. Without a soul in balance, there is no abortion debate. Likewise for the debates over human stem cell research and euthanasia” (p. 16).

To our knowledge, only one study has empirically investigated the relationship between dualist beliefs and bioethics. Richert and Harris (2008) found that people are more ethically opposed to stem-cell research, human cloning, and disconnecting individuals from life support when they perceive those entities to possess souls. They measured soul belief by asking participants questions about the ontology and function of the mind and soul – such as what happens to each after death, and to what extent each can perform a range of spiritual or moral functions, such as contributing to one’s life force and discerning right from wrong. They found that evaluations of the soul, but not the mind, went on to predict responses to the bioethical questions that they posed. That is, people who judged the soul as surviving death, and also as responsible for spiritual and moral functioning, were more likely to oppose these bioethical issues. However, this effect was significantly decreased – almost to the point of non-significance – when accounting for each individual’s level of religiosity, suggesting that commitments to a soul, as measured in their studies, overlapped considerably with religious commitments more broadly.

Richert and Harris’s findings raise a number of important questions. Does belief in an immaterial soul *causally influence* positions on bioethical issues? Or is the causal direction actually the reverse, with bioethical views influencing how people think about the function and ontology of a soul? Or are soul beliefs and bioethical attitudes related through some alternative causal pathway altogether, such as an underlying set of religious commitments? The six studies we report below begin to answer these questions. They also aim not only to isolate religious aspects of soul beliefs from mind-body dualism more generally, but to investigate the roles of related beliefs that have often been conflated with dualism, such as beliefs about reductionism and the scope of science.

Our studies had four central aims. First, we sought to develop a fine-grained measure of dualism that differentiates it from related constructs, such as religiosity and beliefs about science, and that identifies its underlying factor structure. Second, we aimed to confirm the connection between dualism and bioethical judgments posited by Bloom (2004) and Greene (2011), and empirically documented by Richert and Harris (2008). Third, we sought to go beyond this previous work by identifying the specific component(s) of dualism responsible for this connection. Finally, we wished to determine the causal relationships underlying any associations between bioethical judgments and the components of dualism identified by our scale. Studies 1 and 2 speak to the first two aims in that we create and use a novel dualism scale to isolate distinct components of dualism and relate them to bioethical judgments. Studies 3-6 investigate the causal relationships between the two sets of judgments.

Study 1

The aim of Study 1 was to understand and disambiguate the multiple factors involved in dualist attitudes. To do so, we created a novel measurement scale by adapting items from existing scales designed to measure dualism, free will, and determinism (Preston, Ritter & Hepler, 2013; Stanovich, 1989; Nadelhoffer, Shepard, Nahmias, Sripada, & Ross, 2014). We also created original items designed to capture beliefs concerning the metaphysical relationship between the mind and the brain more narrowly, with the aim of potentially dissociating something like “intuitive substance dualism” from the host of related beliefs that have often been measured in conjunction (such as religious beliefs, or those concerning free will).

Our scale also included items concerning the scope of science in explaining the human mind. Past work suggests that mind-body dualism – or belief in an immaterial soul – could be linked with judgments concerning science and reductionism. In particular, Preston, Ritter and Hepler (2013) found that offering participants a neuroscientific explanation (as opposed to a psychological explanation) for the mind decreased their beliefs in an immaterial soul. This suggests a perceived incompatibility between reductive explanations, on the one hand, and a soul, on the other. Findings by Gottlieb and Lombrozo (2017) also suggest, albeit indirectly, that properties that are more central to the mind or soul are more likely to be regarded as falling outside the scope of scientific explanation. Specifically, they found that participants were more likely to judge a phenomenon beyond the scope of science when it was perceived to be uniquely human and tied to a privileged, first-person phenomenology. So, for example, romantic love and altruistic behavior were more likely to be regarded as falling outside the scope of scientific explanation than depth perception or motor control. While these studies did not probe beliefs about souls or dualism, it is plausible that the phenomena judged beyond the scope of science were also those more strongly associated with a mind or soul. A central aim of Study 1 was thus to understand how judgments concerning the scope of science relate to dualist attitudes.

Method: We collected data from 471 participants (248 females, 217 males, 5 who did not identify); mean age = 35 years, $SD = 10$) from Mechanical Turk in exchange for payment. For all studies, we filtered out individuals who had participated in conceptually related studies from our lab group. This also prevented an individual from participating in more than one study reported in this paper. All participants had Mechanical Turk approval ratings greater than 95%, which has been shown to ensure high data quality (Peer, Vosgerau, & Acquisti, 2014). As such, no participants in any of our studies were excluded from analyses, though data from participants who did not complete all parts of the study were not analyzed. Participants answered questions about dualist attitudes and provided demographic information before exiting the study.

We presented participants with 48 items (see Table 3.2; full list can be found on OSF: <https://osf.io/9d7x2/>), some of which were adapted from previous scales measuring dualism, scientific commitments, free will, or determinism (Preston, Ritter & Hepler, 2013; Stanovich, 1989; Nadelhoffer, Shepard, Nahmias, Sripada, & Ross, 2014), and some of which were created for the purposes of this study. We intended the scale to capture a wide range of ideas commonly related to mind-body dualism in the psychological and philosophical literatures, and thus included items related to the following themes: religious commitments to a soul/afterlife (e.g., “God created humans with souls”), limited scope of science (e.g., “Explaining everything that makes us humans in scientific terms in some way decreases the value of life”), mind-brain identity (e.g., “Minds are not the same as brains”), possibility of partial minds (e.g., “Biological organisms are either capable of thought or they aren’t; there is no in-between”), free will (e.g.,

“People always have the ability to do otherwise”), determinism (e.g., “Everything that has ever happened had to happen precisely as it did, given what happened before”), and purity/sacredness (e.g., “The body can be thought of as a temple that houses something with sacred value”). Participants indicated their agreement with each item on a scale from 1 (completely disagree) to 7 (completely agree).

Results: To reduce the dimensionality of our measure, and to extract interpretable and independent dimensions along which individuals varied, we conducted a principle components analysis with varimax rotation. Following Simms and Watson (2007), we excluded items that either 1) loaded at least moderately on more than one component, or 2) loaded together on a non-interpretable factor. We were left with a 32-item scale with five interpretable factors in the rotated solution (see Table 3.1). Table 3.2 presents the factor loadings for each of the 32 items.

Factor name	Theme	Corresponding items	Percent variance explained
Soul/afterlife	Religious commitments to a soul and continued existence after death	1-10	15.60
Limited scope of science	Judgments concerning the limits of science in explaining the mind	11-18	8.18
Free will	The belief that people can consciously will their actions and mental states	24-28	8.09
Determinism	The belief that people's actions are not predetermined	19-23	7.72
Mind-brain identity	Metaphysical judgments that the mind and the brain are not the same	29-32	5.58

Table 3.1 Summary of the Dualism Complex scale and its five components.

Item number	Question	1	2	3	4	5
1	There must be an afterlife of some sort.	0.858				
2	Some nonphysical aspect of us persists after we die.	0.849				
3	Death ends all forms of life forever.*	-0.794				
4	God created humans with souls.	0.804				
5	Nothing in the universe can have nonphysical causes; when people say that God caused something, they are really referring to something that can't be possible.*	-0.709				
6	Every person has a soul.	0.806				
7	What happens to me after I die has nothing to do with how the actions of my life are judged.*	-0.721				
8	If I were asked to sign a piece of paper saying that I would sell my soul to another person, I would feel uneasy doing so.	0.72				
9	The body can be thought of as a temple that houses something with sacred value.	0.674				
10	I cannot wrap my mind around the idea that all my thoughts and feelings are completely gone when I die.	0.515				
11	I think that all aspects of the human mind should be studied scientifically.*		-0.706			
12	I think it would be a good thing if we were someday able to explain all our thoughts and behaviors in scientific terms.*		-0.634			
13	I think that there are some subjects, like love and morality, which should never be studied by scientists.		0.639			
14	There is something troubling about a purely scientific description of mental life.		0.652			
15	Explaining everything that makes us human in scientific terms in some way decreases the value of life		0.59			
16	Knowledge of the mind will forever be beyond the understanding of sciences like physics, neurophysiology, and psychology.		0.59			
17	It is impossible for science to ever have a complete understanding of the mind.		0.59			
18	A complete scientific explanation of the brain will one day be enough to fully describe complex thought.*		-0.508			
19	People's choices and actions must happen precisely the way they do because of the laws of nature and the way things were in the distant past.			0.718		
20	Everything that has ever happened had to happen precisely as it did, given what happened before.			0.717		
21	Every event that has ever occurred, including human decisions and actions, was completely determined by prior events.			0.798		
22	Given the way things were at the Big Bang, there is only one way for everything to happen in the universe after that.			0.67		
23	A supercomputer that could know everything about the way the universe is now could know everything about the way the universe will be in the future.			0.709		
24	People always have the ability to do otherwise.				0.757	
25	People ultimately have complete control over their decisions and actions.				0.841	
26	People always have free will.				0.775	
27	How people's lives unfold is completely up to them.				0.827	
28	People have free will even when their choices are completely limited by external circumstances.				0.737	
29	Minds are not the same as brains.					0.78
30	I believe the mind and brain are the same thing.*					-0.78
31	When people talk about their minds, they are really just talking about what their brains seem to be doing.*					-0.529
32	Minds are in principle independent of bodies.					0.61

Table 3.2 Factor loadings of the 32-item Dualism Complex scale. Factor 1 corresponds to soul/afterlife; factor 2 corresponds to limited scope of science; factor 3 corresponds to free will; factor 4 corresponds to determinism; factor 5 corresponds to mind-brain identity.

Discussion: Study 1 isolated five dissociable components of what has been commonly referred to as psychological or intuitive dualism. Our 32-item scale included items pertaining to a suite of ideas commonly referred to as dualism, but that also incorporate distinct ideas; we therefore refer to the scale as measuring the “Dualism Complex.” Commitments regarding mind-brain identity, which pertain most directly to metaphysical dualism, explained a surprisingly small proportion of variance. Religious commitments to a soul and afterlife beliefs were responsible for the largest percentage of variance, followed by views concerning the limited scope of science in explaining the mind, and then free will and determinism. Having identified these five components, Study 2 examined their relationships to other variables, with a particular focus on bioethical judgments.

Study 2

Study 2 examined the relationship between the dissociable components of our Dualism Complex scale and bioethical attitudes. Doing so had two purposes. First, our study provides a conceptual replication of Richert and Harris (2008). But where they sought to differentiate the soul from the mind, we build on our findings from Study 1 to ask how each component of our Dualism Complex scale relates to bioethical attitudes. In particular, we can examine whether the mind-brain component, which is the most canonically dualist, underlies any association between intuitive dualism and bioethical attitudes. We can also investigate which components – if any – continue to predict bioethical attitudes after controlling for belief in God and political orientation. While soul / afterlife was the component that accounted for most variance in Study 1, Richert and Harris (2008) found that the relationship between belief in a soul and bioethical attitudes was greatly diminished when accounting for religiosity, which suggests that associations between dualism and bioethics could be derivative of a more general set of (religious) commitments. If this is the case, we might expect associations between dualism and bioethics to disappear once we control for belief in god and/or political orientation.

Second, Study 2 sought to test the value of our Dualism Complex measure in two ways. First, if the factor structure found in Study 1 is generalizable and meaningful, then we might expect to find different relationships between the different components and our measures of bioethical attitudes, belief in god, and political orientation. Second, we presented participants with an existing measure of dualism (Stanovich, 1989), which serves as a validation of our measure as well as confirmation that the construct(s) we are elucidating are continuous with those targeted by prior research.

Method: 150 individuals (82 females and 68 males; mean age = 36 years, $SD = 11$) were recruited on Mechanical Turk as in Study 1.

All questionnaires were presented to participants in a random order, ending with demographic information. We measured dualist attitudes using the 32-item Dualism Complex Scale developed in Study 1. In addition, participants completed a single question measuring political orientation (“In general, how politically liberal or conservative do you consider yourself?” Inbar, Pizarro, & Bloom, 2009), and a four-item measure of belief in god (e.g., “When I am in trouble, I find myself wanting to ask God for help;” Shenhav, Rand, & Greene, 2012). We also included an existing 27-item dualism scale (Stanovich, 1989) to examine the

relationship between our scale and an existing measure.

We measured bioethical judgments using items adapted from the Gallup Poll. For each issue (abortion, physician-assisted suicide, cloning animals, cloning humans, and medical research using stem cells obtained from human embryos), participants were asked, “Regardless of whether or not you think it should be legal, for each one, please indicate whether you personally believe that in general it is morally acceptable or morally wrong.” Ratings ranged from 1 (morally acceptable) to 7 (morally wrong).

Results: We first investigated how the Dualism Complex Scale relates to an existing measure of dualism (Stanovich, 1989). The correlation was high ($r = .75, p < .001$); this helps establish continuity between our measure and past work that has aimed to study intuitive dualism. Responses to the Dualism Complex scale were also highly correlated with belief in god ($r = .71, p < .001$) and with political conservatism ($r = .48, p < .001$): the more a participant believed in God or identified as conservative, the more likely that participant was to endorse the attitudes assessed by the Dualism Complex Scale.

Of primary interest was whether the five components identified in Study 1 related to bioethical judgments *differentially*. We computed “bioethical scores” by averaging ratings of the five bioethical issues. Lower ratings corresponded to the view that the queried activities are generally morally acceptable, and higher ratings to the view that they are not. This average rating was significantly correlated with the Dualism Complex scale as a whole ($r = .50, p < .001$), as well as several of its subcomponents: participants were more likely to find these activities morally unacceptable the more strongly they scored on soul/afterlife ($r = .57$), on limited scope of science ($r = .42$), on free will ($r = .19$), and on mind-brain identity ($r = .17$), all p 's $< .05$. Determinism was only marginally related to bioethical scores, and trended in the opposite direction ($r = -.15, p = .062$). Each of these correlations became insignificant when controlling for belief in god and political conservatism (p 's $> .10$), with the exception of limited scope of science ($r = .17, p = .018$).

We also entered all five components in a multiple regression to identify the unique variance accounted for by each. We found significant effects of soul/afterlife ($b = .53, \beta = .49, p < .001$), scope of science ($b = .34, \beta = .25, p < .001$), and mind-brain identity ($b = -.19, \beta = -.17, p = .028$). Interestingly, mind-brain identity was *negatively* related to bioethical judgments when the other dualist components were accounted for (i.e., people who believe the mind and brain are independent were more likely to rate bioethical issues as morally acceptable). Since both dualism and bioethical judgments were strongly related to belief in god and to political conservatism, we fit an additional model in which we treated belief in god and political conservatism as covariates. In this model, the two covariates absorbed much of the variance accounted for by soul/afterlife, and only limited scope of science remained a significant predictor of bioethical attitudes ($b = .19, \beta = .26, p = .011$).

Discussion. Study 2 revealed three key findings. First, the Dualism Complex scale tracks an individual difference closely related to the construct targeted by Stanovich's (1989) existing measure of dualism, as the two measures were highly correlated. However, our scale includes additional structure in the form of five subcomponents that relate to our other measures in unique and interpretable ways. This suggests the value of our measure as a tool for further research. Second, we found that the mind-brain identity component was a weak predictor of bioethical attitudes, and that it did not predict significant variance beyond belief in god and political conservatism. This is potentially surprising given that it is arguably the most conceptually central to dualism, at least as understood within philosophy, and as typically discussed within

psychology.

Third, we found that among the five components of dualist attitudes isolated in Study 1, limited scope of science shared a particularly robust relationship with bioethical judgments. Of the three components that independently predicted bioethical judgments – soul/afterlife, mind-brain identity, and limited scope of science – only limited scope of science remained significant after statistically controlling for belief in god and political conservatism. While scope of science beliefs have received some attention (Gottlieb & Lombrozo, 2017), these findings are the first to relate them to bioethical attitudes.

Why might scope of science judgments predict bioethical attitudes? In the studies that follow, we unpack this relationship. Specifically, Studies 3-6 aim to disentangle possible causal structures that could explain the correlation between scope of science commitments and bioethical judgments. In Study 3, we test the hypothesis that scope of science commitments causally influence bioethical judgments. In Study 4, we test the hypothesis that bioethical judgments causally influence scope of science commitments. Lastly, Studies 5 and 6 test the hypothesis that bioethical judgments and scope of science commitments share a common set of determinants.

Study 3

Study 3 tested the hypothesis that scope of science commitments causally influence bioethical judgments. We manipulated participants' scope of science commitments using a passage arguing either that science could explain many aspects of the human mind (broad scope of science condition) or few aspects of the human mind (narrow scope of science condition). Following this passage, all participants read about a hypothetical mammal that was discovered on a remote island in the Pacific Ocean, and indicated how morally permissible it would be to conduct stem-cell research on this mammal and to clone this mammal. We chose to ask participants about a hypothetical new species, as opposed to humans, as we expected people to have at least somewhat defined antecedent positions about the moral permissibility of cloning or performing stem-cell research with humans, thus making it less likely that we could experimentally manipulate their bioethical positions. We also chose to focus on just cloning and stem-cell research as these were the two issues (of the five) included in Study 2 that were amenable to describing this hypothetical new species; specifically, it seemed implausible to ask people about performing abortions or physician-assisted suicide on a non-human animal. We expected that participants in the broad scope of science condition would consider cloning and stem-cell research more morally permissible than those in the narrow scope of science condition, or in a control condition that read about a neutral topic.

Method: 318 individuals (145 females, 172 males, and 1 who did not specify; mean age = 37 years, $SD = 13$) were recruited from Amazon Mechanical Turk as in Studies 1-2. Participants were randomly assigned to read one of three texts: one arguing that science could fully explain the human mind (broad scope of science), one arguing that science could never fully explain the human mind (narrow scope of science), and one that provided unrelated information about the role of science in cooking (control). In all three conditions, participants were asked to read a three-paragraph passage, which was framed as an essay that a student wrote about “the most important thing” they learned from an introductory course. Those in the two scope of science conditions read an essay called “What Neuroscience Can Teach Us.” These essays drew their wording from the scope of science component of our Dualism Complex scale to include statements like, “The most important thing that I learned this semester is that all aspects of the

human mind can and should be studied scientifically,” “This course taught me that it would be very valuable if we were someday able to explain all our thoughts and behaviors in strictly scientific terms,” and, “A scientific understanding of human experience is both achievable and desirable.” The narrow scope of science passage used parallel wording to communicate the idea that science cannot and should not fully explain the human mind and behavior. Lastly, participants in the control condition read an essay titled “What Cooking Can Teach Us,” which included no content about the scope of science with regard to the human mind. The passages were roughly matched in length (ranging from 215-239 words) and are available at OSF (<https://osf.io/9d7x2/>).

Immediately after reading one of these three essays, participants were told to suppose that scientists have discovered a new mammal on a remote island in the Pacific that looks similar to a sheep, but is a distinct species. They were asked two questions about this mammal: 1) “To what extent do you think it would be morally permissible for scientists to clone this mammal for research purposes?” and 2) “To what extent do you think it would be morally permissible for scientists to conduct research using embryonic stem cells obtained from this species?” Participants indicated their judgment on a scale from 1 (not at all morally permissible) to 7 (completely morally permissible). Participants also indicated their agreement with 12 manipulation check items taken from the Dualism Complex scale created in Study 1. Four questions measured scope of science commitments, and the remaining questions corresponded to other components (mind-brain identity, soul/afterlife, free will, and determinism). All questionnaire materials can be accessed on OSF (<https://osf.io/9d7x2/>). Participants indicated their agreement with each item on a scale from 1 (completely disagree) to 7 (completely agree).

Results: We first tested the efficacy of our manipulation by testing for an effect of condition on scope of science commitments. Participants in the narrow scope of science condition were most likely to agree with statements saying that science cannot fully explain the human mind ($M = 4.50$, $SD = 1.72$), followed by those in the control condition ($M = 3.88$, $SD = 1.57$), and finally by those in the broad scope of science condition ($M = 3.75$, $SD = 1.57$). A one-way ANOVA revealed a significant effect of condition, $F(2,315) = 6.49$, $p = .002$, and independent-samples t-tests revealed significant differences between the narrow scope of science condition and both the neutral condition, $t(208) = 2.76$, $p = .006$, and the broad scope of science condition, $t(207) = 3.31$, $p = .001$, which did not differ from each other, $t(211) = .62$, $p = .535$. There were no group differences in ratings of mind-brain identity, $F(2,315) = 2.06$, $p = .130$, soul/afterlife, $F(2,315) = .34$, $p = .710$, or free will, $F(2,315) = .80$, $p = .452$, though there was a marginal effect on determinism, $F(2,315) = 3.00$, $p = .051$. This suggests that our manipulation was effective, and moreover supports the independence of the components of the Dualism Complex scale, given that an intervention on one component had minimal effects on others.

We next tested for effects of experimental condition on the two bioethical judgments. There were no effects of condition on the moral permissibility of cloning, $F(2,315) = .28$, $p = .758$, nor on the moral permissibility of stem-cell research, $F(2,315) = .32$, $p = .729$. Interestingly, however, variation in scope of science beliefs (as assessed by the manipulation check) was correlated with bioethical judgments: people with narrower scope of science commitments were more opposed to cloning ($b = -.50$, $\beta = -.40$, $p < .001$), and to stem-cell research ($b = -.45$, $\beta = -.39$, $p < .001$). So although condition significantly predicted scope of science commitments, and scope of science commitments significantly predicted bioethical judgments, we failed to find a significant relationship between condition and bioethical judgments.

Discussion: Study 3 tested the hypothesis that scope of science commitments influence bioethical judgments. We failed to find evidence for a causal relationship of this form. Before concluding that no such relationship exists, however, a few cautionary remarks are in order. One is that this conclusion rests on a null result, and so could reflect inadequate statistical power. Reassuringly, we did find a significant effect of our passages on the scope of science manipulation check items, suggesting that the passages were effective and that we had sufficient power to detect an effect on these measures. That said, it could still be that the study was underpowered with respect to the effect of condition on bioethical judgments, if such an effect is small. With our sample size, we had an 80 percent chance of detecting a medium to large effect size.

Another possibility is that our manipulation or dependent variables were somehow flawed. With respect to this concern, the significant correlation between the bioethical dependent measures and the scope of science manipulation check items is potentially relevant. Were the problem some mismatch between what we manipulated and what we assessed – because one involved humans and the other sheep, for example – we would not necessarily expect this correlation to emerge.

Although concerns about power and our measures are somewhat mitigated by these considerations, we felt it prudent to replicate our null result with different materials. We conducted a number of additional studies with slight variation to the passages and dependent variables; these studies similarly failed to find an effect of a scope of science manipulation on bioethical attitudes, despite there being an effect of experimental prime on scope of science judgments. We therefore tentatively conclude that the causal structure “scope of science → bioethical judgments” is inadequate, and go on to test alternatives in Studies 4-5.

Study 4

In Study 4, we tested the possibility that bioethical judgments causally influence scope of science commitments. Given previous theorizing on the relationship between dualism and bioethics (Bloom, 2004; Greene, 2011), we expected to find that experimentally manipulating scope of science commitments would causally influence bioethical judgments. However, since Study 3 (and its conceptual replications) failed to find evidence for this relationship, we asked whether the latter causally influences the former. To test this hypothesis, we manipulated participants’ views on animal cloning by presenting them with a text passage explaining either the benefits or drawbacks of a pet cloning service. Participants then answered questions about their scope of science commitments.

Method: 200 individuals (87 females and 113 males; mean age = 32 years, $SD = 10$) were recruited from Amazon Mechanical Turk as in Studies 1-3. Participants were randomly assigned to one of two conditions, one of which advocated for cloning, and one of which argued against cloning. Participants in both conditions were told that they would read an essay that another student wrote about the topic of pet cloning. The essay advocating for cloning was titled, “If your dog is about to die, why not clone it?” It described a South Korean researcher who clones pets, and included statements like, “I believe that these services are invaluable and worth any price tag. Pet cloning is an important and valuable advance, and any pet owner will benefit from the availability of this service.” The essay advocating against cloning was titled, “If your dog is about to die, you should not clone it,” and stated, “It is foolish to believe that these services would benefit society. Pet cloning is a moral offense, and I strongly believe that any pet owner can understand why we should not support this endeavor.” The essays were matched for

approximate length (188-200 words), and are available at OSF (<https://osf.io/9d7x2/>).

Immediately following the essay, participants were asked to write, in their own words, why pet cloning is either positive or negative, depending on their experimental condition. They were asked to write for at least two minutes, at which point they were able to continue to the next page of the study. We included this task to reinforce and strengthen the intended experimental manipulation.

Participants then answered six questions taken from our Dualism Complex scale. Four questions measured scope of science commitments, and two questions measured views on mind-brain identity. These items were identical to the ones used as a manipulation check in Study 3, but items from the other dimensions were not included. Participants indicated their agreement with each statement on a scale from 1 (completely disagree) to 7 (completely agree). Finally, participants responded to three manipulation check questions about pet cloning. These questions measured how much they supported or opposed the idea of paying to have a pet cloned, how much they would support the development of such a service in the United States, and how uncomfortable they would be with the idea that a dog they knew could be cloned. Participants indicated their judgment using a sliding bar that ranged from zero to 100; high scores indicated support for, and comfort with, the idea of pet cloning

Results: We first tested the efficacy of our manipulation by comparing scores on the three questions about pet cloning. As expected, those who read the essay advocating for pet cloning were more in favor of the technology ($M = 44.44$, $SD = 28.75$) than those in the other condition ($M = 30.75$, $SD = 23.53$), $t(184) = 3.70$, $p < .001$. There was also a significant correlation between people's judgments of pet cloning and their ratings on the scope of science items ($r = .34$, $p < .001$). However, we failed to find an effect of condition on scope of science commitments, as reflected in participants' averaged ratings for the four scope of science items, $t(198) = .49$, $p = .625$. We also failed to find an effect on averaged ratings for the two mind-brain identity items, $t(198) = 1.24$, $p = .217$.

Discussion: Study 4 failed to find support for the hypothesis that bioethical attitudes and scope of science commitments are correlated because the former exert a causal influence on the latter. Given the null result, the same concerns that arose for Study 3 hold here. It is therefore relevant to point out that the manipulation was effective (as reflected by a significant effect of condition on the manipulation check items), and that scope of science beliefs were once again significantly correlated with the measure of bioethical attitudes. We did run one additional conceptual replication of this study, and similarly found a null result. As a consequence, we moved on in Study 5 to test a final possibility: that bioethical attitudes and scope of science beliefs are correlated because they share common determinants.

Study 5

If scope of science commitments and bioethical judgments share a common set of determinants, what might those determinants be? In past research (Gottlieb & Lombrozo, 2017), we found that people are more resistant to scientific explanations for some mental phenomena than for others. The phenomena most strongly deemed to fall beyond the scope of science were those for which people reported privileged introspective access (e.g., feeling empathy), over which they felt that they have conscious will (e.g., acting morally), and that were perceived as making humans exceptional compared to other species (e.g., appreciating music). Since we know that these factors influence scope of science commitments, and also that scope of science commitments relate to bioethical judgments, in Study 5 we asked whether these factors also

influence bioethical judgments. If so, they could explain why bioethical judgments and scope of science judgments are related.

Method: 100 participants (62 females and 38 males; mean age = 34 years, $SD = 10$ years) were recruited on Amazon Mechanical Turk as in Studies 1-4. We followed a within-subjects design with three conditions. In each experimental condition, participants answered questions about stem-cell research and the scope of science, but they did so with respect to a fictional mammal with relevant human-like traits (human-like condition), with respect to a fictional mammal that lacked those traits (non human-like condition), or with respect to humans (human condition). The order of the three conditions was randomized across participants.

In the first two conditions, participants were introduced to a hypothetical new mammal that scientists had discovered on a remote island in the Pacific. This mammal was described as being similar to sheep, but in fact a distinct species. The human-like condition then described this mammal as possessing the three qualities we previously found to be related to scope of science commitments (the ability to introspect, the ability to exert conscious will, and as possessing uniquely human qualities; Gottlieb & Lombrozo, 2017), and the non human-like condition described the mammal as lacking these qualities. Below are excerpts from the relevant conditions:

Human-like mammal: Scientists have learned quite a bit about this mammal's mind and behavior. They have shown that this mammal actually acts quite similarly to humans in the sense that it has the ability to introspect.

Non human-like mammal: Scientists have learned quite a bit about this mammal's mind and behavior. They have shown that this mammal actually doesn't have the ability to introspect the way that humans do.

Participants were then asked: "Suppose that scientists are interested in using stem cells obtained from this species' embryos for research purposes. To what extent do you think it is morally permissible for them to conduct research using these embryonic stem cells?" Participants indicated their response on a scale from 1 (not at all morally permissible) to 7 (completely morally permissible). They were also asked the same four questions about the scope of science that were used in previous studies, but this time the wording was adapted to refer to this species: 1) "I think that all aspects of this mammal's mind should be studied scientifically" (reverse coded), 2) "I think that there would be something troubling about a purely scientific description of this mammal's mental life," 3) "It would be impossible for science to ever have a complete understanding of this mammal's mind," and 4) "Knowledge of this mammal's mind will forever be beyond scientific understanding." Participants responded on a scale from 1 ("completely disagree") to 7 ("completely agree"). These questions were identically worded in the two conditions. In the third (human) condition, participants read no introductory text, and were asked the moral permissibility of conducting research using human embryonic stem cells. They also answered the same four scope of science questions, but this time the questions referred to the human mind, and thus were identical to the ones used in previous studies.

Results: We first evaluated whether the experimental manipulation of properties affected bioethical judgments. We conducted an ANOVA with condition as a within-subjects variable and the average rating for the moral permissibility of stem-cell research as the dependent variable, and found a significant effect of condition, $F(2,198) = 7.71, p < .001$. Participants rated stem cell research as most morally permissible when it involved the mammal that lacked the three human-like qualities (i.e., the non human-like mammal condition; $M = 5.28, SD = 1.60$; see Figure 3.1), followed by the human condition ($M = 5.05, SD = 1.83$), and finally by the mammal

that possessed the three human-like qualities (i.e., the human-like mammal condition; $M = 4.77$, $SD = 1.84$). Paired t-tests revealed a significant difference between the two mammal conditions, $t(99) = -3.91$, $p < .001$, the human-like mammal condition compared with the human condition, $t(99) = -2.19$, $p = .031$, and a marginally significant difference between non human-like mammal and human conditions, $t(99) = 1.75$, $p = .084$.

We next evaluated whether the experimental manipulation of properties affected scope of science judgments. A within-subjects ANOVA with average scope of science judgments as the dependent variable revealed a significant effect of condition, $F(2,198) = 5.21$, $p = .006$.

Participants evaluated science as most able to fully explain the mind of the non human-like mammal ($M = 3.53$, $SD = 1.48$; see Figure 3.1), followed by the human mind ($M = 3.82$, $SD = 1.50$), and finally the mind of the human-like mammal ($M = 3.83$, $SD = 1.45$). Paired samples t-tests revealed significant differences between the two mammal conditions, $t(99) = -2.53$, $p = .013$, and the non human-like mammal condition compared with the human condition, $t(99) = -2.54$, $p = .013$, but no difference between the human and human-like mammal conditions, $t(99) = .06$, $p = .950$.

Discussion. The results of Study 5 suggest that bioethical judgments and scope of science judgments share a common set of determinants. Specifically, we manipulated several characteristics of a fictional species to make its mind more or less human-like, and we found that participants were most likely to judge that science could one day explain aspects of the mind when that species lacked human qualities: the ability to introspect, the ability to exert conscious will, and as possessing qualities thought to be uniquely human. Participants also judged it as more morally permissible to conduct stem-cell research using cells from this fictional species compared to the other two conditions.

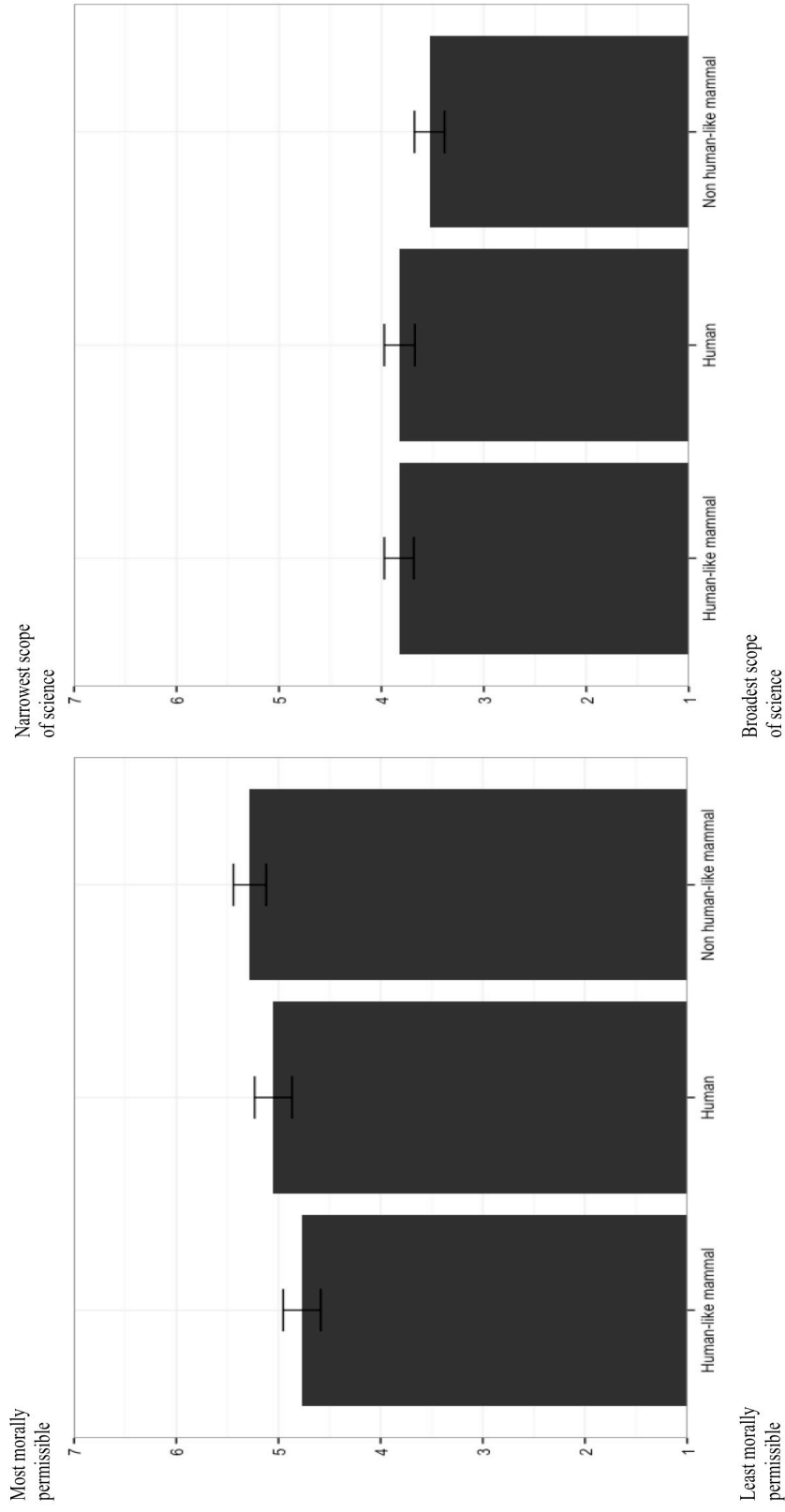


Figure 3.1 Results from Study 5. Average bioethical judgment by condition (left); higher scores indicate that performing stem-cell research on that species is judged more morally permissible. Average scope of science judgments by condition (right); higher scores indicate narrower scope of science judgments (e.g., the belief that science can explain fewer aspects of that species' mind). Error bars represent one standard error above and below the mean.

Study 6

The results of Study 5 suggest that certain aspects of the mind – the ability to introspect, the ability to exert conscious will, and qualities perceived as unique to humans – can causally influence both scope of science commitments and also bioethical judgments. Participants judged embryonic stem-cell research as most morally permissible when performed on mammals described as lacking those three mental qualities, and also judged that science could best explain the minds of such mammals.

Study 6 extended these findings in three ways. First, we aimed to conceptually replicate the findings of Study 5. Second, we tested the generalizability of the effect by extending our findings to a new bioethical judgment. We again presented participants with questions that referred to either (non-human) mammals or humans, but this time probed them about the moral permissibility of cloning. Finally, Study 6 independently examined the three factors we found to influence both scope of science commitments and bioethical judgments in Study 5 (the ability to introspect, the ability to exert conscious will, and the capacity to possess uniquely human qualities).

Method: 300 individuals (182 females, 117 males, and 1 who did not specify; mean age = 35 years, $SD = 11$) were recruited from Amazon Mechanical Turk as in Studies 1-5. Participants were asked a question about cloning, but we again varied the information they were shown prior to this question, this time using a between-subjects design. In the human condition, participants were simply asked to consider humans. In the control condition, they are asked to consider a hypothetical new mammal discovered on a remote island in the Pacific, as in Study 5, but were presented with no additional information about the mammal. In the remaining six conditions, participants also read about this hypothetical new mammal, but were additionally told that it either *possessed* or *lacked* one of the three properties of interest: the ability to introspect (introspection + or -), the ability to exert conscious will (conscious will + or -), or putatively uniquely human traits, such as the ability to experience complex emotions or romantic love (uniquely human + or -). For example, participants in the introspection (+) condition read the following:

Suppose that scientists have discovered a new mammal on a remote island in the Pacific. This mammal is similar to a sheep in the sense that it looks like a sheep, but it is actually a distinct species.

Scientists have shown that this mammal actually acts quite similarly to humans in the sense that it has the ability to introspect. For example, when having a certain experience - like an emotion or desire - it can know what it is experiencing by examining its own feelings and reflecting on them. This means that it is privileged to a kind of information about what it is experiencing that the scientist observing it can't know in the same way.

Participants in the conscious will (-) condition instead read that: "Scientists have shown that, unlike humans, this mammal actually can't consciously or deliberately influence when, why, or

how many of their actions or behaviors will occur. For example, it does not have the ability to consciously will itself to resist a temptation or to make a choice among several options” (see OSF for the full set of stimuli: <https://osf.io/9d7x2/>).

All participants answered the following question about cloning, which referred either to humans (in the human condition) or the mammal (in the remaining seven conditions): “Suppose that scientists are interested in cloning this animal [humans] for research purposes. To what extent do you think it is morally permissible for them to clone this mammal [humans]?” Participants indicated their response on a scale from 1 (“not at all morally permissible”) to 7 (“completely morally permissible”).

Results: Participants rated cloning humans as least morally permissible ($M = 2.92$, $SD = 2.03$), and cloning the mammal that lacked the ability to exert conscious will as most morally permissible ($M = 5.11$, $SD = 1.79$; see Figure 3.2). In between these two extremes were the following conditions (in order from least to most morally permissible): the mammal described as possessing uniquely human qualities ($M = 3.64$, $SD = 2.19$), the ability to introspect ($M = 3.84$, $SD = 1.73$), and the ability to exert conscious will ($M = 4.14$, $SD = 2.17$); the mammal described as lacking uniquely human qualities ($M = 4.44$, $SD = 2.09$), lacking the ability to introspect ($M = 4.63$, $SD = 1.98$), and finally the control condition for which we provided no additional information about this mammal ($M = 4.76$, $SD = 1.58$). A one-way ANOVA revealed a significant effect of condition: $F(7,292) = 4.89$, $p < .001$.

We compared means between pairs of conditions that were of particular interest. First, we compared the conditions that described the hypothetical mammal as possessing or lacking a certain quality. Participants rated cloning as significantly more morally permissible for the mammal that lacked conscious will than for the mammal that possessed conscious will, $t(71) = -2.08$, $p = .041$. There were no significant differences between the conditions that described the mammal as possessing or lacking the ability to introspect, $t(74) = -1.85$, $p = .068$, or as possessing or lacking uniquely human qualities, $t(73) = -1.62$, $p = .11$.

We also compared all ratings to those of our two reference points: humans and the control condition. All conditions, with the exception of the one that described the mammal as having uniquely human qualities, were significantly different from the human condition. The control condition was significantly different from those that described the mammal as having the ability to introspect, $t(74) = -2.42$, $p = .018$, as having uniquely human qualities, $t(75) = -2.57$, $p = .012$, and also the condition that referred to humans, $t(74) = -4.41$, $p < .001$.

Discussion. Study 6 demonstrated that framing a mammal as having (or lacking) one of three properties known to relate to people’s scope of science commitments influenced the degree to which participants believed it would be morally permissible to clone the being in question. While our results do not clearly point to one factor as driving the effect observed in Study 5 (all seemed to have an effect), we saw that the presence versus absence of conscious will might be especially powerful.

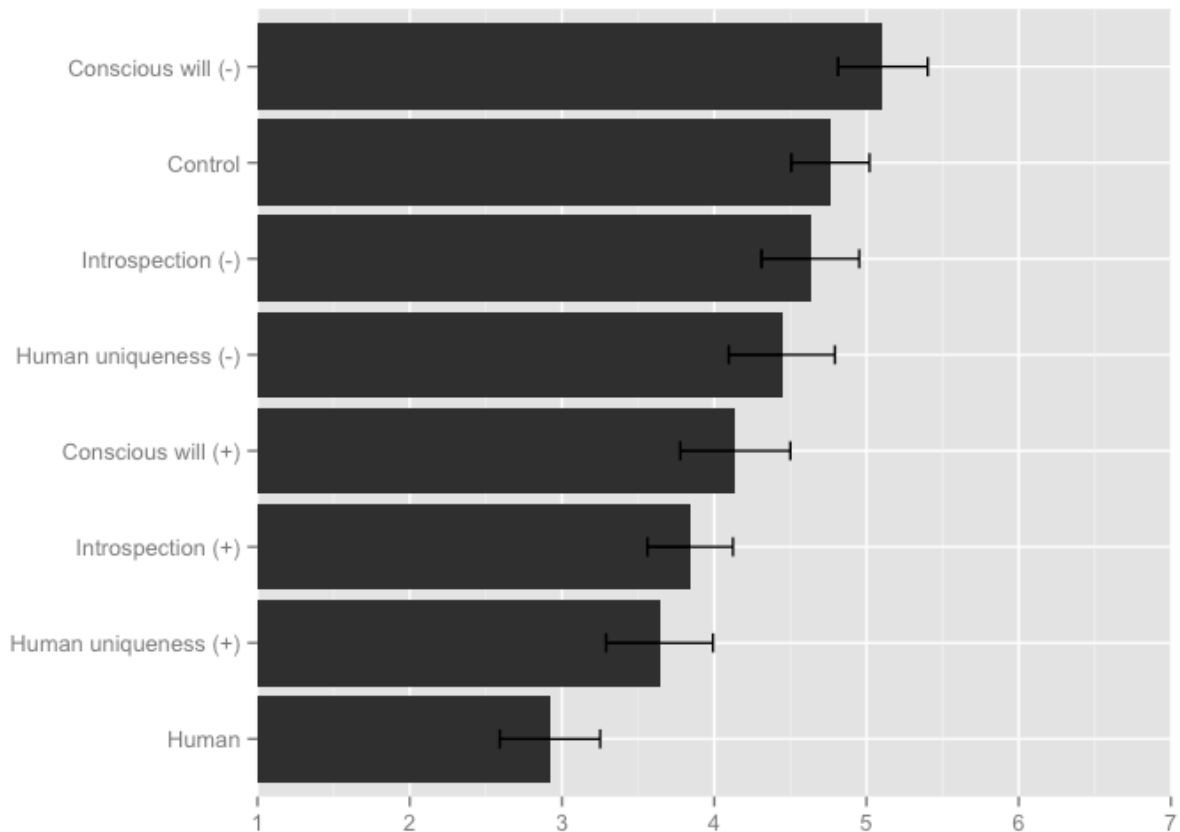


Figure 3.2 Mean ratings of the moral permissibility of cloning by experimental condition (Study 6). Higher scores indicate judgments of greater moral permissibility. Error bars represent one standard error above and below the mean.

Discussion

Previous theoretical and empirical work has pointed to a connection between dualist commitments and bioethical attitudes (Bloom, 2004; Greene, 2011; Richert & Harris, 2008), but the causal structure relating the two has remained unclear. In the present set of studies, we sought to better characterize this relationship and, in doing so, isolated distinct components of intuitive mind-body dualism. In Study 1, we found that attitudes concerning mind-brain identity – arguably the most “dualist” in the metaphysical sense – accounted for a relatively small proportion of variance in people’s views compared to religious commitments to a soul and the afterlife, and also views on the appropriate scope of science in explaining the mind. We found in Study 2 that only scope of science commitments, and not other aspects of dualism, predicted bioethical judgments independently of religiosity (and conservatism) more broadly. In Studies 3 and 4 we failed to find a straightforward causal relationship between scope of science commitments and bioethical attitudes, and data from Studies 5 and 6 instead suggest that the two share a common set of determinants. More specifically, we found that when the mind of an organism is perceived as having certain qualities – such as the ability to exert conscious will or the ability to introspect on its own mental activity – people tend to perceive it as being beyond

the scope of scientific explanation, and they also tend to judge some interventions on it, such as cloning or stem-cell research, less morally permissible.

While the causal story we found evidence for is not straightforward, it raises important questions regarding the relationship between psychological dualism and views about science and reductionism. Much has been written on the topic, but little work (e.g., Preston, Ritter & Hepler, 2013) has empirically investigated how attitudes toward science, and neuroscience in particular, are causally related to, or even wholly distinct from, psychological dualism. We ultimately found that more basic commitments about mental phenomena – such as one’s ability to introspect or exert conscious will – causally influences judgments about whether science can explain mental life and behavior, and also bioethical judgments.

These results do, however, raise yet more questions about the relationship between scope of science commitments and bioethical attitudes. In Studies 5 and 6, we intervened on several aspects of a species’ mind, and showed that such a manipulation causally influenced both scope of science commitments and bioethical attitudes. Nonetheless, it still remains unclear how these relationships manifest in the real world. It might be that people who increasingly attribute certain qualities to humans – such as the ability to introspect and exert conscious will – are more likely to judge these bioethical issues as innocuous, since the entity in question is in less possession of what is typically considered a mind or essential spirit. However, more work is needed to fully understand the causal structure at play, and to isolate other determinants that might also account for the relationship between scope of science commitments and bioethical attitudes.

Also notable is that, in Study 2, judgments concerning mind-brain identity were only weakly related to dualist attitudes. Further, they shared a weaker relationship with bioethical attitudes than commitments to a soul or commitments concerning the scope of science, and this relationship was even diminished to non-significance when accounting for religiosity. Our studies suggest that intuitive dualism as studied so far may have weak ties to metaphysical dualism as defined in philosophy. However, a suite of related beliefs – most notably, attitudes toward science and reductionism – play an important role in human judgments and behaviors.

Nonetheless, our findings clarify that there does indeed exist a relationship between dualism and bioethical attitudes – and one that exists independent of religiosity or political orientation – even if it is an aspect of “dualism” that departs from traditional definitions in both the philosophical and psychological literature. We have provided a new tool for studying intuitive dualism and the suite of associated commitments with which it is often packaged. Moving forward, we urge the field to adopt a narrower scope when referring to dualism in the more metaphysical sense, and to more clearly isolate the aspect(s) of interest when attempting to study dualism more broadly. We do not yet know why scope of science commitments appear so central to dualist attitudes; it might reflect an important aspect of dualism’s factor structure, or it might more simply reflect a greater tendency for an individual’s judgments to vary on these judgments as opposed to ones concerning metaphysics. Nonetheless, our findings do suggest that studying judgments about the scope of science is itself a part of the puzzle of better understanding intuitive dualism and its downstream consequences on our thoughts and behaviors.

Chapter 4. Choosing when we cannot know: First-person priority in transformative choice

In a typical day, we face hundreds of decisions. Most of these are mundane and low-stakes, but some decisions – such as the decision to take a mind-altering drug, to change one’s religious identity, or to have a child of one’s own – are potentially transformative. How do people evaluate what the outcomes of such decisions will be like, and how do they ultimately decide which option to pursue?

“Transformative” decisions seem to pose a special puzzle for accounts of rational choice (Paul, 2014). In making a more mundane decision – such as buying a new car – a rational agent can compare the expected utility of each option (Briggs, 2017). But if an experience (such as having a child) is truly transformative, then the pre-choice individual cannot readily evaluate the outcome of the transformative choice: she is not on a position to know what the experience will be like, nor does she have the same values as her post-decisional self. More precisely, the experience is potentially transformative *epistemically*, in the sense that she gains new experiential knowledge as a consequence of the decision, as well as *personally*, in the sense that it radically changes her core preferences and identity.

To some extent, the puzzle of transformative choice is a puzzle for all choice: people are often bad at predicting their affective responses to future experiences (even mundane ones; Wilson & Gilbert, 2013), and philosophers disagree about whether the distinction between transformative and non-transformative experience is truly a difference in kind or merely degree (Paul, 2014; Chang, 2015). Nonetheless, by highlighting puzzles of choice so forcefully, transformative decisions offer a unique window onto human decision making, including the epistemic and normative commitments that underlie it. How do people think they can *know* what an experience will be like, and how do they think they *ought* to make the decision? Do these beliefs vary as a function of how transformative the decision is taken to be?

One way in which transformative and non-transformative decision-making could differ is in the role of first-person sources of information, such as reflection and imagination, versus second and third-person sources of information, such as testimony or scientific studies. For transformative decisions, a first-person perspective is especially compromised: by definition, the pre-choice individual cannot know what the experience will be like (Jackson, 1986; Nagel, 1974). Testimony and science bypass this limitation by relying on information from individuals who have already undergone the relevant transformation. A decision-maker can solicit the testimony of parents, for example, or she can consult scientific studies comparing parents and non-parents in terms of their happiness or well-being. But ought she make a transformative decision on the basis of such sources?

Reliance on testimony or science might solve an epistemic problem, but they seem to introduce another: these sources can’t reveal what an experience will be like *for her*. Moreover, relying on testimony or science to make a transformative decision, such as having a child, might threaten people’s sense of authentic decision-making or personal agency (Paul, 2014). Finally, people could resist the idea that decisions involving phenomena with a rich first-person phenomenology, such as parental or romantic love, even fall within the scope of third-personal, scientific investigation (Gottlieb & Lombrozo, 2017). So while imagination might be especially limited when it comes to *knowing* what a transformative experience will be like, people might nonetheless favor it over these alternatives as a basis for *making* the transformative decision that they perceive as right *for them*. This motivates a striking prediction: that while first-personal sources of information will be more compromised when it comes to knowing what a

transformative experience would be like (relative to a non-transformative experience), imagination – relative to testimony and science – will be favored when it comes to *making* such decisions.

Across three studies, we investigate the extent to which people think imagination, testimony, and science allow them to know what an experience would be like, and the extent to which they think each source can appropriately guide them to making a decision about whether or not to engage in that experience themselves. We consider both transformative and non-transformative decisions, with the expectation that anticipated epistemic and personal transformation will affect the relative contribution of each source of information. Specifically, in Studies 1a and 1b, we test the predictions that people will recognize their limited ability to imagine what transformative (as opposed to non-transformative) experiences are like, but will nonetheless privilege imagination – especially compared to testimony or science – when judging how transformative decisions ought to be made. In Study 2, we extend these findings by investigating people’s judgments on a wider range of decisions, and by isolating the contributions of epistemic versus personal transformation. Finally, in Study 3, we validate our findings from hypothetical decisions by having participants report the process by which they made actual decisions.

Study 1a

In Study 1a, we asked participants to evaluate the extent to which they could know what transformative experiences (such as having a child) and non-transformative experiences (such as having a knee replacement) are like by consulting imagination, testimony, and science. We predicted that imagination would be judged a less effective source of knowledge for transformative decisions than for non-transformative decisions.

Method: Two-hundred and three individuals (93 males, 106 females, and 4 who did not specify; mean age = 35 years, $SD = 12$) were recruited from Amazon Mechanical Turk in exchange for payment. For all studies, we filtered out individuals who had participated in conceptually related studies from our lab group. This also prevented an individual from participating in more than one study reported in this paper. All participants had Mechanical Turk approval ratings greater than 95%, which has been shown to ensure high data quality (Peer, Vosgerau, & Acquisti, 2014). As such, no participants in any of our studies were excluded from analyses, though data from participants who did not complete all parts of the study were not analyzed. We collected pilot data from 100 individuals, and the sample size for Studies 1a and 1b was determined by doubling this number. The sample size and analysis plan were preregistered on OSF (<https://osf.io/dx5k3/>).

Study 1a followed a fully within-subjects design with two factors: experience type (2: transformative, non-transformative) and information source (3: science, testimony, imagination). All participants were asked to consider eight hypothetical decisions, four of which are typically considered transformative (having a child, undergoing sex reassignment surgery, taking a mind-altering drug, and undergoing a religious conversion), and four of which are considered non-transformative (taking out a life insurance plan, having a knee replacement, taking a long-term cholesterol medication, and transferring to a different school). For each scenario, individuals were asked to indicate the degree to which they could know what that experience would be like by relying on scientific information, testimonial evidence, or their own imaginations. Below are sample questions for the item *having a child* (italicized labels were not presented to participants):

Science: To what extent could you know what it is like to have a child by reading scientific studies that describe people who have children on a variety of dimensions, such as their health, satisfaction with their choices, and/or well-being?

Testimony: To what extent could you know what it is like to have a child by talking to other people who have children, and listening to what they have to say about their experiences?

Imagination: To what extent could you know what it is like to have a child by thinking and imagining what it would be like to have a child?

Participants indicated their answers on a scale ranging from 1 (not at all) to 7 (completely). The order of scenarios – as well as the order of questions within each scenario – was randomized for each participant.

Results: For each participant, we created six scores by collapsing across the four transformative or non-transformative decisions, computing an average corresponding to each information source. We then conducted a repeated-measures ANOVA with experience type and information source as within-subjects factors. This revealed a main effect of experience type, $F(1,202) = 108.80, p < .001, \eta^2 = .04$: participants gave lower ratings for transformative experiences than for non-transformative experiences ($M = 3.79, SD = 1.56$ and $M = 4.28, SD = 1.50$), indicating that, across information sources, they judged it less possible to know what a transformative experience would be like relative to a non-transformative experience. We also found a main effect of information source, $F(2,404) = 275.88, p < .001, \eta^2 = .34$. Participants thought they could best know what an experience would be like through testimony ($M = 5.02, SD = 1.24$), followed by science ($M = 4.20, SD = 1.27$) and imagination ($M = 2.89, SD = 1.26$); paired samples t-tests revealed that all pairs were significantly different from one another (all p 's $< .001$).

Importantly, we also found a significant interaction between experience type and information source, $F(2,404) = 21.45, p < .001, \eta^2 = .01$ (see Figure 4.1). Although participants thought it was more possible to know what a non-transformative experience would be like, regardless of information source, this difference was most pronounced in the domain of science (mean difference = $-.74, t(194) = 11.88, p < .001$), followed by testimony (mean difference = $-.45, t(194) = 7.51, p < .001$) and imagination (mean difference = $-.27, t(194) = 4.31, p < .001$). So while imagination was indeed perceived to be a worse guide to knowing what a transformative decision (versus a non-transformative decision) would be like, as predicted, it was also the case that second and third-person sources of information were especially compromised by the transformative nature of a choice.

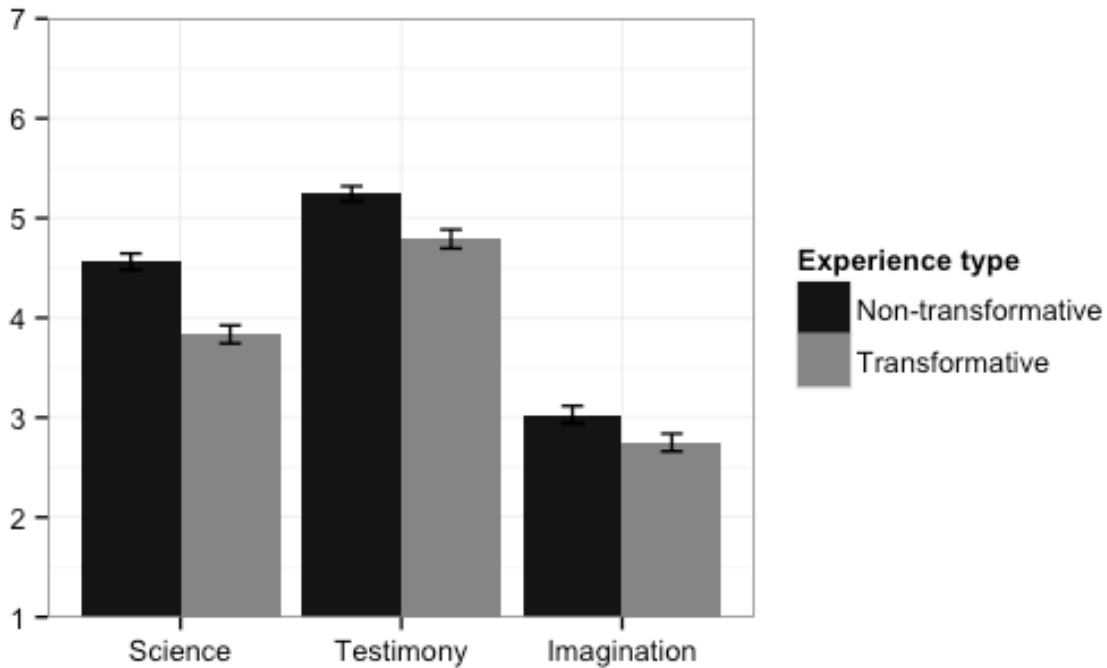


Figure 4.1 Average “know” judgments from Study 1a.

Study 1b

In Study 1b, we asked participants to evaluate the extent to which they ought to rely on imagination, testimony, and science to make the best decision regarding the same transformative and non-transformative experiences used in Study 1a. We predicted that imagination would be judged a good guide to decision-making, especially for transformative decisions.

Method: One hundred and ninety-four individuals (89 males, 103 females, and 2 who did not specify; mean age = 35 years, $SD = 10$) were recruited from Amazon Mechanical Turk in exchange for payment.

The procedure was identical to that of Study 1a, except that participants made judgments about how they *ought to decide* whether or not to engage in transformative and non-transformative experiences. They were asked (for example): “Suppose that you are deciding whether or not to have a child. In order to make the ultimate decision that is best for you, how important is it to do the following?” We then asked them to rate from 1 (not at all) to 7 (very much) the following three questions:

Science: Read scientific studies that compare individuals with and without children on a variety of dimensions, such as personal happiness.

Testimony: Talk to individuals who both have and do not have children, and listen to what they have to say about their experiences.

Imagination: Consider what it would mean to have a child, and also what it would mean to not have a child, by thinking and imagining what these two scenarios would be like.

The order of scenarios, and questions within scenarios, were both randomized across participants.

Results: We again created six averages corresponding to the ratings across the four transformative or non-transformative decisions for each information source. We then conducted a repeated-measures ANOVA with experience type (2: transformative, non-transformative) and information source (3: science, testimony, imagination) as within-subjects factors. This revealed a significant main effect of experience type, $F(1,193) = 12.49, p < .001, \eta^2 = .01$: participants thought they should consult information sources to a greater extent for non-transformative experiences ($M = 5.56, SD = 1.11$) than for transformative experiences ($M = 5.35, SD = 1.30$). We also found a significant main effect of information source, $F(2,386) = 73.22, p < .001, \eta^2 = .10$: participants thought they ought to rely most on their imaginations ($M = 5.85, SD = 1.08$), followed by testimony ($M = 5.57, SD = 1.18$) and science ($M = 4.95, SD = 1.22$). Paired samples t-tests revealed that all pairs were significantly different from one another (all p 's $< .001$).

Once again, we also found a significant interaction between experience type and information source, $F(2,192) = 39.49, p < .001, \eta^2 = .02$ (see Figure 4.2). Participants thought they ought to rely on science and testimony more for non-transformative (versus transformative) decisions (mean difference = .65, $t(193) = 7.98, p < .001$ for science, and mean difference = .19, $t(193) = 2.33, p = .021$ for testimony). However, this pattern was reversed when it came to imagination, for which participants indicated that they ought to rely on their imaginations more when deciding whether or not to engage in a transformative (as opposed to non-transformative) experience (mean difference = .21, $t(193) = 2.65, p = .009$). This confirms our prediction that imagination would be judged a good guide to decision-making (especially for transformative decisions), and additionally suggests that second and third-person sources were penalized in making a transformative choice.

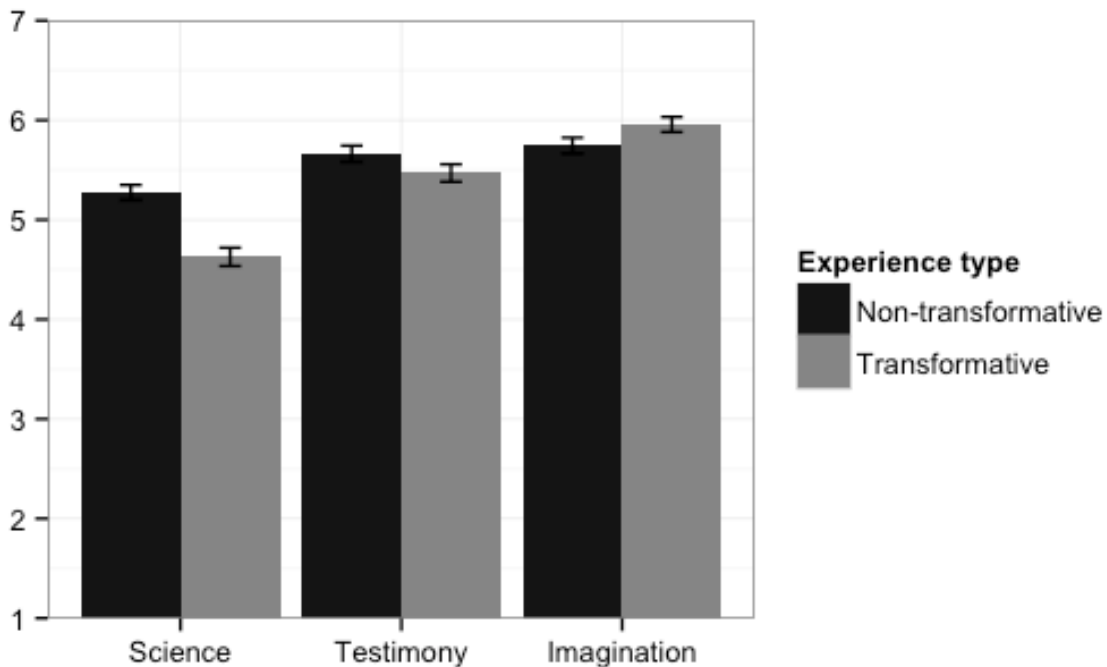


Figure 4.2 Average “decide” judgments from Study 1b.

Discussion: Studies 1a and 1b reveal an interesting mismatch between participants' beliefs about how they could know what an experience is like, on the one hand, and how they thought a decision about it ought to be made, on the other. Participants rated testimony as the best source of information for knowing and imagination as the worst, with science falling in between. Yet when it came to judgments about *deciding*, they rated imagination superior to testimony, with science rated lowest of all. Though these patterns held for both transformative and non-transformative decisions, they were most pronounced for the former: for transformative decisions, imagination was judged especially compromised for knowing, but especially important for deciding.

Study 2

Study 2 aimed to extend the results from Studies 1a-1b in three ways. First, we tested a much larger and more diverse set of experiences (over 40), which allowed us to verify the generality of our results and to conduct analyses across items rather than across individuals. Second, we asked participants to separately evaluate the extent to which each experience would change them *epistemically* and *personally* in order to dissociate these two dimensions of transformative decisions. Finally, given that transformative decisions often have greater consequences than non-transformative decisions, we controlled for the relative “stakes” of each decision to ensure that this factor was not driving the effects found in Studies 1a and 1b.

Method: Five-hundred and forty-two individuals (289 males, 252 females, and 1 who did not specify; mean age = 36 years, $SD = 11$) were recruited from Amazon Mechanical Turk in exchange for payment. The sample size was determined by using a similar number of stimulus items and individuals to Gottlieb and Lombrozo (2017), which employed a similar across-item analysis.

All participants were presented with a list of 41 experiences. We designed this set of experiences to range from not at all transformative (e.g., choosing a new computer to buy) to highly transformative (e.g., deciding to have a child), and also to vary on epistemic and personal transformativeness (all stimuli can be found at OSF: <https://osf.io/qsjty/>). Participants rated all 41 experiences along *one* of nine dimensions, making “dimension” a between-subjects factor.

Six of the nine dimensions were those used in Studies 1a and 1b. Individuals assigned to the three “know” dimensions were asked about the extent to which they could know what that experience would be like through either science, testimony, or imagination, using wording like that in Study 1a. Individuals assigned to one of the three “decide” dimensions were asked how much they ought to rely on science, testimony, or imagination when deciding whether or not to engage in that experience, with wording like that used in Study 1b.

Two of the dimensions were designed to assess how transformative each experience was perceived to be, with the aim of dissociating epistemic and personal transformation. The epistemic transformativeness question asked (for example): “To what extent does having a child teach you something you could not have learned without having that experience itself?” The personal transformativeness question asked (for example): “To what extent does having a child change who you are in some deep and personally fundamental way, for example, by altering things like your core personal preferences, your desires, and your point of view?” Participants answered on a scale from 1 (not at all) to 7 (very much).

The final dimension asked about the stakes, or relative importance, of making a decision about each experience, and it was included as a control. Participants indicated, on a scale from 1

(not at all) to 7 (very much), how strongly they agreed with a statement such as: “Deciding to have a child is a choice that matters, has important consequences, and should be taken seriously.”

Results: Between 53 and 73 participants provided data for each dimension (with an average of 65 per dimension). We computed each item’s average score on each dimension and then conducted analyses across items (as opposed to across individuals), with nine data points for each experience.

We first examined data on the three “know” dimensions to investigate how participants thought they could know what a diverse range of experiences would be like. A repeated-measures ANOVA with information source as a within-subjects factor revealed a significant effect of information source, $F(2,80) = 25.48, p < .001, \eta^2 = .06$: participants thought they could best know what an experience would be like by relying on testimony ($M = 3.97, SD = .73$), followed by imagination ($M = 3.82, SD = .97$), and science ($M = 3.48, SD = .69$). Paired samples t-tests revealed that all pairs were significantly different from one another (p ’s $< .001$ for testimony and science and imagination and science, and $p = .045$ for testimony and imagination). This mirrors the findings from Study 1a in that testimony dominated, though in the present study, imagination was rated higher than science.

To investigate how the transformativeness of a given experience – either epistemic or personal – related to these judgments, we fit three separate regression models using epistemic transformativeness, personal transformativeness, and stakes to predict the degree to which participants thought they could know about that experience through science, testimony, or imagination, respectively. Epistemic and personal transformativeness were the predictors of interest, with stakes included to control for the fact that transformative decisions are often higher stakes than non-transformative decisions. For all three models, we found a significant (negative) effect of epistemic transformativeness ($b = -.58, p < .001$ for science, $b = -.58, p < .001$ for testimony, and $b = -.64, p = .008$ for imagination), but no effects of personal transformativeness or stakes. In other words, the more epistemically transformative the experience, the less participants thought that they could know what it would be like through any information source. These results also mirror those from Study 1a, but go beyond them in identifying *epistemic* transformation as the factor responsible for the main effect of experience type found in that study.

We next analyzed the three “decide” dimensions to investigate how participants thought they ought to decide whether or not to engage in each experience. We performed a repeated-measures ANOVA with information source as a within-subjects factor, and we again found a significant effect of information source, $F(2,80) = 26.90, p < .001, \eta^2 = .20$. Participants thought they ought to rely most heavily on their imaginations ($M = 5.06, SD = .76$), followed by testimony ($M = 4.82, SD = .47$) and science ($M = 4.21, SD = .90$). Paired samples t-tests revealed that all pairs were significantly different from one another (p ’s $< .001$ for imagination and testimony and testimony and science, and $p = .023$ for imagination and testimony). This mirrors the findings from Study 1b in that imagination dominated the other non first-person sources.

We next fit three separate regression models using epistemic transformativeness, personal transformativeness, and stakes to predict these three “decide” ratings. For the model with science as the outcome variable, we found significant effects of epistemic transformativeness ($b = .43, p = .038$), personal transformativeness ($b = -.48, p = .023$) and stakes ($b = .67, p < .001$). In other words, participants thought they should rely on science to a greater extent for decisions that were more epistemically transformative and higher stakes, but that they should rely on science *less* for decisions that were more personally transformative. The model using testimony as the outcome

variable revealed similar effects for personal transformation ($b = -.26, p = .023$) and for stakes ($b = .35, p < .001$). Finally, for the model using imagination as the outcome variable, we found only a significant effect of stakes ($b = .67, p < .001$). These findings go beyond Study 1b in identifying *personal* transformation as the dimension responsible for the more circumscribed role for science in making transformative decisions. Based on Study 1b, we would also have expected to find a positive effect of epistemic or personal transformation on imagination, but this was not found.

Discussion: Across a wide range of hypothetical decisions, we replicated key findings from Studies 1a and 1b. Notably, testimony was rated most highly when it came to *knowing* (as in Study 1a), but imagination was rated most highly when it came to *deciding* (as in Study 1b). Study 2 went beyond these results in isolating the factors due to epistemic versus personal transformation. While epistemic transformation drove the lower ratings for how an experience can be known (regardless of source), personal transformation was responsible for participants' unwillingness to value testimony and science – the two non first-personal sources – more highly when deciding whether or not to pursue a transformative choice. Finally, these results were obtained while controlling for stakes.

Study 3

In Study 3, we sought to go beyond normative judgments about hypothetical decisions to reported behavior about real decisions. To do so, we asked participants to tell us about the degree to which they consulted different sources of information when making transformative and non-transformative decisions in their own lives.

Method: 447 individuals (220 males, 220 females, and 4 who did not specify; mean age = 36 years, $SD = 11$) were recruited from Amazon Mechanical Turk in exchange for payment. We followed a mixed-design with experience type (transformative, non-transformative) as a between-subjects factor, and information source (science, testimony, imagination) as a within-subjects factor.

Participants were randomly assigned to answer questions about either a transformative or non-transformative decision they once had to make, and were asked to select one decision from a list of four. Those in the transformative condition chose from: deciding whether or not to have a child, deciding whether or not to end a relationship, deciding whether or not to get married, and deciding whether or not to change their religious beliefs. Those in the non-transformative condition chose from: deciding whether or not to undergo laser eye surgery, deciding whether or not to transfer to a new job, deciding whether or not to buy a new car, and deciding whether or not to legally change their name. Seven of these eight options were taken from the stimuli used in Study 2; we examined average ratings for transformativeness (defined as the product of epistemic and personal transformativeness), and chose 3-4 decisions that we thought participants were likely to have experienced. The three transformative decisions were rated among the 20 most transformative, and the four non-transformative decisions were rated among the 11 least transformative. The eighth option, ending a relationship, was not included in Study 2, but was added as an option to increase the likelihood that every participant assigned to the transformative conditions would have experienced at least one of the four provided options. Across both conditions, 27 individuals indicated that they had never made any of these decisions; thus, data from only 420 participants were included in analyses (206 in the transformative condition, and 214 in the non-transformative condition).

After selecting a decision, participants then answered questions specific to the decision they selected. They were asked, in random order, three questions about how they made the decision, two questions about the transformativeness of that decision, and one question about stakes. The transformativeness and stakes questions were worded identically to those in Study 2. We amended the wording of the decision questions slightly to more closely approximate real-world decision-making, for example:

Science: When making this decision, to what extent did you rely on scientific evidence or statistical information to consider and compare the two options (have a child versus not having a child)? Examples could include reading articles or media coverage comparing the health or happiness of people who do and do not have children, or reading statistics about people who made each choice.

Testimony: When making this decision, to what extent did you rely on the testimony of other people to consider and compare the two options (having a child versus not having a child)? Examples could include talking with friends who do or do not have children to hear about their feelings and experiences, or reading memoirs or personal anecdotes.

Imagination: When making this decision, to what extent did you rely on your imagination and reasoning to consider and compare the two options (have a child versus not having a child)? Examples could include imagining your future self as a parent, or thinking about what it would feel like to have or not to have children at various points in your life;”

Participants indicated their answer on a scale ranging from 1 (not at all) to 7 (very much). They also indicated whether or not they ultimately decided to undergo the experience in question (e.g., chose to have a child).

Results: As a manipulation check, we compared transformativeness ratings across experience types using independent-samples t-tests. Transformative decisions were indeed rated more epistemically transformative (means = 5.01 and 4.14, respectively, $t(418) = 4.87, p < .001$) and more personally transformative (means = 4.77 and 3.31, respectively, $t(418) = 7.84, p < .001$) than non-transformative decisions. The two did not differ in stakes (means = 6.35 and 6.22 in the transformative and non-transformative conditions, respectively, $t(418) = 1.29, p = .20$).

Overall, participants indicated ultimately pursuing the queried option (e.g., having a child or undergoing laser eye surgery) in 80% of cases (175 for transformative, 159 for non-transformative). As the findings did not differ as a function of the participants’ choice, we collapsed the data across the variable (three-way interaction between participant choice, experience type, and information source: $F(2,828) = 2.01, p = .13$).

To evaluate the relationship between experience type and reported information source, we conducted a 2x3 mixed-effects ANOVA with experience type as a between-subjects factor and information source as a within-subjects factor. Replicating Study 1b, we found a significant main effect of experience type, $F(1,416) = 7.59, p = .006, \eta^2 = .01$: participants reported consulting information sources to a greater extent for non-transformative experiences ($M = 4.47, SD = 1.86$) than for transformative experiences ($M = 4.22, SD = 2.05$). We also found a significant main effect of information source, $F(2,832) = 44.52, p < .001, \eta^2 = .07$: participants reported relying most heavily on their imaginations ($M = 5.53, SD = 1.39$), followed by testimony ($M = 4.00, SD = 1.79$) and science ($M = 3.50, SD = 2.03$). Paired samples t-tests revealed that all pairs were significantly different from one another (all p 's < .02).

As expected, we once again found a significant interaction between experience type and information source, $F(2,832) = 9.78, p < .001, \eta^2 = .02$ (see Figure 4.3). The observed pattern closely mirrors that of Study 1b: participants reported relying on science and testimony more heavily when making non-transformative (versus transformative) decisions, but relied on imagination to a greater extent when making transformative (versus non-transformative) decisions. Independent samples t-tests revealed a significant effect of experience type within each information source ($t(418) = 3.20, p = .001$ for science, $t(418) = 2.83, p = .005$ for testimony, and $t(418) = 2.56, p = .011$ for imagination).

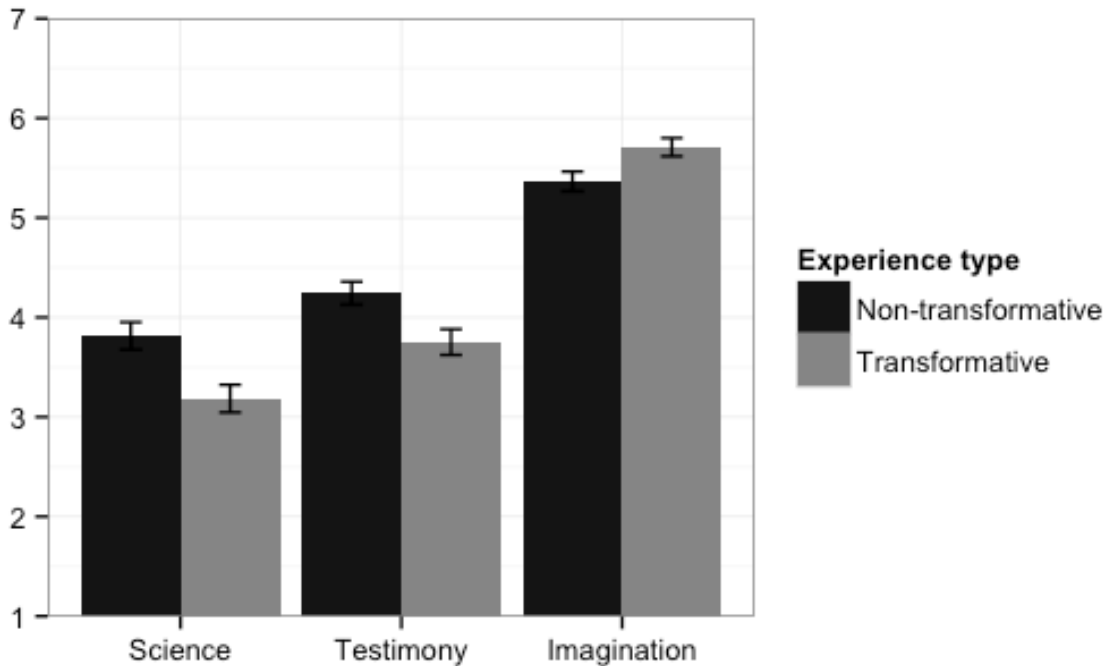


Figure 4.3 The degree to which individuals in Study 3 reported relying on science, testimony, and imagination when making transformative or non-transformative decisions.

Discussion

Across three studies, we find that people value testimony most highly when it comes to knowing what the outcome of a decision would be like, but favor their own imagination and reflection when it comes to making the best decision *for them*. This is the case for both transformative and non-transformative decisions (Studies 1- 2), with the latter effect observed in both hypothetical (Studies 1b, Study 2) and actual (Study 3) decision-making.

Beyond this striking mismatch between people's epistemic judgments (of how they can know) and their normative judgments (of how they ought to decide), we found moderating effects of perceived personal and epistemic transformation. The more epistemically transformative the experience, the less any information source could reveal what it's like. But the more personally transformative the experience, the greater the priority for first-person (relative to second and third-person) sources. This manifested as a penalty for science and testimony in making decisions about transformative (versus non-transformative) experiences, and in Studies 1b and 3, as an advantage for imagination.

Why might we observe a mismatch between judgments about *knowing* versus *deciding*? While the former judgment was specific to the experience of pursuing the transformative option (e.g., knowing what it's like to have a child), the latter judgment required a comparison of both options (e.g., not having a child versus having a child) along multiple dimensions, some of which could extend beyond the experience itself (to consider, for example, the environmental impact of having a child). To explain the priority of imagination when it comes to deciding, however, it seems that more is required: why would a comparative evaluation along multiple dimensions result in first-person priority? To answer this question, the moderating effect of perceived transformation provides some clues. One possibility is that people are resistant to the idea that others' experiences or scientific generalizations apply to *them*, and that this resistance is stronger the more transformative the experience. Indeed, there's evidence that people regard phenomena with a rich first-person phenomenology, such as love at first sight, as falling beyond the scope of science (Gottlieb & Lombrozo, 2017), and many of our transformative experiences shared this characteristic. Another possibility is that deferring to others for deeply personal decisions is regarded as an inauthentic decision-making procedure (Campbell, 2015; Mogensen, 2017), or that it signals a lack of true understanding on the decision-maker's part (Paul, forthcoming). These possibilities could apply to both transformative and non-transformative decisions, albeit to different degrees, and raise important questions for future research.

Querying people's *beliefs* about decision making cannot, on its own, address the rationality of their decision-making process, or the quality of their decisions. Nonetheless, the correspondence between the findings from Study 1b regarding hypothetical decisions, and those from Study 3 regarding actual decisions, suggests that people's normative commitments regarding the value of science, testimony, and imagination guide their actual decision-making. It's an open question whether these behaviors support better decisions, or influence other downstream consequences, such as post-choice satisfaction and regret.

While our studies focus on the issue of transformative choice, our findings reveal epistemic and normative commitments that apply more broadly. We must decide, on a daily basis, whether or not to engage in a range of experiences, typically without knowing precisely what those experiences will be like. It might seem natural to assume that, in making everyday choices, people would believe it best to rely on those sources of information that they deem most likely to reveal what that experience will be like. Our results, however, suggest that this is not the case.

Chapter 5. Discussion

I began this dissertation with the case of Mary the color scientist who is said to know all the physical facts about color and human color vision. She has spent her entire life studying the science of color, and has a full scientific understanding of the topic. But Mary has never actually seen color. When she steps out of her black and white chamber for the first time, she seems to learn something new. Her scientific knowledge does not prepare her to know *what it is like* to see red, or blue, or yellow.

Frank Jackson's (1986) thought experiment elicits strong intuitions about the limits of science and scientific explanation; though science can provide a thorough understanding of many topics, it may be more limited in giving us information about the subjective experience of mental states and various aspects of the human experience. For some types of experiences, people seem to give special status to their own introspective knowledge over other types of information. Across a series of studies, I have investigated people's folk epistemic commitments on topics that converge on two central themes: 1) do some types of knowledge fall beyond the scope of science? And 2) can science serve as a good guide to decision-making?

Across these three sets of studies, I have shown that people tend to privilege their own first-person knowledge over science – an inherently third-person source of information – when it comes to certain aspects of the human mind and experience. In Chapters 2 and 3, this manifested in people judging science as unable to ever fully explain aspects of the mind to which they feel that they have privileged first-person access. In Chapter 4, this manifested as people privileging their own first-person knowledge when making decisions, especially those that are regarded as personally transformative. These findings advance the literature on folk epistemology, but also demonstrate that epistemic commitments have important consequences for moral judgments and how people go about making decisions in their everyday lives.

In this final chapter, I will outline key findings from each set of studies and suggest directions for future research. Finally, I will discuss the implications of these results for the real limits of scientific explanation and understanding.

Do some types of knowledge fall beyond the scope of science?

Epistemic commitments

Chapter 2 investigated intuitive commitments concerning the limits of science in explaining the human mind. I found that people generally thought that science could least explain mental phenomena to which they have privileged first-person access (e.g., feeling empathetic), those that contribute to making humans exceptional compared to other species (e.g., the ability to appreciate a beautiful sunset), and those that can be consciously willed (e.g., acting altruistically).

Across these studies, I looked at both epistemic commitments about the *limits* of science, and also people's non-epistemic judgments about what science should or should not study. People often communicate a discomfort with the idea that science could explain topics typically regarded as sacred – things like religious belief, romantic love, or moral judgment. Interestingly, I found that epistemic commitments about the limits of science are closely linked to these non-epistemic judgments, which may be value-based or affective in nature. I asked people the extent to which they would be uncomfortable with the idea that science could explain a range of topics related to the human mind, and again found that people were uncomfortable with science explaining mental phenomena to which they have privileged first-person access, those that

contribute to making humans exceptional and above other species, and those that can be consciously willed.

These results are among the first to empirically demonstrate why some topics are considered more amenable to scientific investigation than others, but echo a sentiment that is not entirely new. One nice example comes from a 2012 commentary published in *Nature*. The author and scientist Daniel Sarewitz wrote about his experience visiting the Angkor temples in Cambodia, and compared it to the discovery of the Higgs boson. He described the powerful sense of mystery and transcendence elicited by the temples, alongside sensing the enormity of the universe that evaded comprehension. In contrast, he wrote of the Higgs boson: “Science is supposed to challenge this type of quasi-mystical subjective experience, to provide an antidote to it” (Sarewitz, 2012, p. 431). He goes on to explain that religion – or any type of transcendent experience, like that of the Angkor temples – can provide what he considers an authentic encounter with something greater, whereas the Higgs is only “an incomprehensible abstraction, a partial solution to an extraordinarily rarified and perhaps always-incomplete intellectual puzzle.” The upshot of his argument is that there are some things that science cannot, and perhaps should not, aim to provide as far as knowledge goes; these include personal encounters with the unknown, or insight into the mystery of existence.

That said, there do remain a number of unanswered questions worthy of future research. I queried people’s judgments of science and scientific explanation, but the present set of studies does not rule out the possibility that people think that topics like love or morality simply cannot be explained or understood, be it by science or some other discipline altogether (e.g., religious studies, literature). It is thus possible that these results reflect a more general tendency to think that other entities or institutions – or even other individuals – cannot explain a given mental phenomenon, and future work is necessary to test this hypothesis.

Relatedly, there remain open questions concerning scientific explanations themselves. It is important to note that I asked people about the potential for a complete scientific explanation, but never actually provided them with a hypothetical explanation. Could it be that once offered, scientific explanations would actually be accepted, or perhaps even welcome? There is evidence that people actually prefer explanations for psychological phenomena that appeal to neuroscience, as opposed to just psychology. Specifically, past research has documented what is known as the “reductive allure” effect: people prefer explanations at lower levels (e.g., ones that appeal to neuroscience) compared with explanations at higher levels (e.g., that appeal only to psychology), even when the lower-level explanations do not include any additional explanatory content.

I suspect that when it comes to evaluating explanations for the types of phenomena that tend to fall beyond the scope of science – things like romantic love or religious experience – the allure of increasingly reductionist explanations will be offset by the allure of intuitive dualism and narrow scope of science commitments. Given people’s intuitively dualist tendencies, it is possible that, despite the “reductive allure,” people would be more uncomfortable with explanations for love, for example, as those explanations become increasingly reductive. More specifically, I predict that people would be more uncomfortable with a biological explanation than a neuroscientific one, and more uncomfortable with a neuroscientific explanation than a psychological one. It is less clear how people might evaluate epistemic judgments about what science could *possibly* explain; it may be, for example, that people judge psychological explanations as most possible given that they also find these as easiest to understand or generate

on one's own (Keil, Lockhart, & Schlegel, 2010), despite the fact that they would find them more uncomfortable.

Intuitive dualism

The studies presented in Chapter 2 are among the first to investigate people's judgments concerning the limits of science in explaining the mind, but were also partially predicted by previous research on intuitive dualism (e.g., Bloom, 2004). Humans are proposed to intuitively view the world in terms of material bodies and immaterial minds or souls, and often have a difficult time reconciling scientific evidence that our thoughts and behaviors are rooted in the material brain. For this reason, I expected that aspects of the mind having to do with a soul or essential spirit – in other words, those to which we can introspect, or can consciously will – would be considered beyond the scope of science, and so the results of Chapter 2 contribute an important theoretical connection between mind-body dualism and scope of science commitments.

Past work on intuitive dualism has, to some extent, already involved discussions of scope of science commitments (e.g., Preston, Ritter, & Hepler, 2013), but dualism has also been understood of as referring to a suite of related concepts (e.g., mind-brain identity, free will, determinism, and religious commitments to a soul). In Chapter 3, I sought to clarify the construct of intuitive dualism by understanding its factor structure and dissociable elements. Interestingly, I found that scope of science commitments account for a fairly large amount of variance in dualist attitudes, despite the fact that, in both psychology and philosophy, dualism has been defined primarily as tracking attitudes about the relationship between the mind and brain (Descartes, 1641; Robinson, 2017). This result is noteworthy as it suggests that intuitive dualism as studied so far may have weak ties to metaphysical dualism as defined in philosophy.

Past theoretical and empirical work on intuitive dualism (Bloom, 2004; Greene, 2011; Richert & Harris, 2008) also suggests that dualist commitments are responsible, in part, for opposition to bioethical issues such as abortion or cloning. In Study 2, I used the dissociable elements of intuitive mind-body dualism to predict bioethical attitudes and found, surprisingly, that only scope of science commitments – but not other aspects of dualism – related to these moral judgments independently of political and religious orientation. This result goes beyond previous work to demonstrate that it is a precise notion of dualism, specifically scope of science commitments, which is responsible for this relationship.

Little work has studied people's moral judgments about bioethics in particular, but previous research on disgust does provide some context within which we can interpret this result. Disgust is often linked with purity-based moral values (Horberg, Oveis, Keltner & Cohen, 2009) and conservative voting behavior (Inbar, Pizarro, Iyer, & Haidt, 2012). This link between disgust and socio-moral attitudes is thought to be strongest for sanctity issues like abortion (Inbar, Pizarro, & Bloom, 2009; Kumar, 2018). Accordingly, disgust has even been referred to as the “body and soul” emotion as it is theorized to protect not only the physical body from contaminants, but also the immaterial body from immorality or impurity (Rozin, Haidt & McCauley, 1999). To my knowledge, no empirical work – correlational or experimental – has demonstrated a relationship between dualism and disgust, but individual differences in dualism, and perhaps scope of science commitments in particular, might arise from more basic sensitivities to disgust, or the other way around. It is commonly observed that politically conservative individuals tend to be more opposed to science and scientific advance, especially in a domain like genetics that generates bioethical debate. Despite this fact, it is not particularly

well understood why science and bioethics is such a politically divisive issue, but future work involving dualism and scope of science commitments may help take a step toward untangling this mystery.

Taken together, Chapters 2 and 3 document important epistemic commitments regarding the limits of science when it comes to explaining the human mind. As suggested by Jackson's (1986) thought experiment involving Mary the color scientist, people do judge certain types of mental phenomena or experiences as beyond the scope of scientific explanation, and these are the types of phenomena for which people privilege their own introspective access and first-person knowledge. These commitments have important consequences for bioethical decision-making, and likely have downstream effects for a range of other topics, including the public acceptance and dissemination of science, and neuroscience in particular.

Can science serve as a good guide to decision-making?

The case of Mary the color scientist raises questions regarding the extent to which relying on physical facts can and should be a reliable guide to decision-making. When Mary is afforded freedom from her chamber, it might seem obvious that she should take this opportunity to step out into the sunshine. But if she cannot anticipate what the subjective experience of seeing color will be like, are we able to say that she is making a rationally informed decision? This is an extreme example, but a similar dilemma arises for deaf individuals who are contemplating receiving cochlear implants. Without knowing what it is like to have this new sensory ability, these individuals are unable to assign subjective value to what the new experience would be like *for them*, despite the fact that hearing is regarded as a normatively good capability.

The studies I presented in Chapter 4 focus on two central questions: to what extent do people think that science (and also imagination and testimony) can allow them to know what new experiences will be like? And to what extent do they think that science (and also imagination and testimony) can guide them in deciding whether or not to engage in these new experiences? I specifically asked whether the answers to these two questions vary for transformative and non-transformative experiences.

Transformative experiences presented an interesting lens through which I was able to approach these questions. In virtue of these experiences being epistemically transformative, individuals are – before having the actual experiences themselves – unable to anticipate the subjective experience of what these outcomes would be like for their future selves. Despite the fact that philosophers have discussed this puzzle of transformative choice (Paul, 2014), it was an open empirical question of how people would generally think they ought to – and actually do – go about deciding whether or not to engage in these types of experiences.

In Study 1a, I asked people how they could know what a range of experiences would be like by consulting scientific evidence, the testimony of others, or their own imaginations. In Study 1b, I asked them the extent to which they should rely on these three sources when deciding whether or not to engage in those experiences. These sources of information map onto third, second, and first-person perspectives, respectively. In line with what Paul (2014) would predict, I found an interesting mismatch between these two sets of judgments: people judged their own imaginations – a sort of first-person source – as most relevant when deciding, despite the fact that they also judged it difficult to imagine what these experiences would be like. This effect was most dramatic for transformative experiences: people thought that imagination was the worst way to know what these experiences would be like, but nonetheless judged imagination as the most appropriate basis for a decision.

People's judgments of science were also particularly noteworthy. They generally considered science especially bad for both knowing what transformative experiences would be like, and also deciding whether or not to engage in them. In Study 2, I specifically found that this unwillingness to privilege scientific information was related to the degree to which an experience was considered personally transformative. In other words, people privileged their own first-person knowledge for these types of cases.

Interestingly, personally transformative experiences often have many of the qualities that I found in Chapter 2 to affect scope of science judgments. More precisely, transformative experiences – which are often considered highly individual, sacred experiences – have a distinctive phenomenology associated with them, are accessible only to the experiencer herself, and contribute to making humans exceptional compared to other species. Canonical personally transformative experiences include falling in love, transcendent spiritual or religious experiences, or feeling awed by a beautiful landscape. These are among the same types of experiences that I found to be regarded as beyond the scope of science in Chapter 2.

In Study 3, I asked people to report how they actually went about deciding whether or not to engage in transformative and non-transformative decisions. I observed a strikingly similar pattern to people's normative judgments about how judgments *ought* to be made: people were, overall, most likely to privilege their own imaginations when deciding, and this was especially true for transformative choices. Perhaps even more striking, however, was that people reported a particularly low reliance on science for *all* type of decisions; this is something I would have expected for transformative choices, but it is more surprising that people devalued science even for the non-transformative cases, which included typically mundane things, like buying a new car or deciding whether or not to undergo laser eye surgery. Taken together, I found that this first-person privileging is most salient for transformative choices, but may actually reflect a more general tendency to devalue scientific evidence in the decision-making process.

Why do people demonstrate such a privileging of imagination (a first-person source of information) over science (a third-person source of information) when it comes to deciding? Much future work is necessary to understand this phenomenon. One possibility is that people tend to think that scientific or testimonial evidence is not generalizable to *them*, specifically. This would be especially true for scientific evidence, which tends to report measures of central tendency. People rarely fall at the mean or median of a given parameter, so it is possible that they are intuitively committed to the idea that a statistic is a poor indicator of how they, as individuals, should think or behave.

A second possibility is that people find that deferring to others – whether that be deferring to the testimony of other individuals, or consulting empirical sources – constitutes an inauthentic decision-making procedure (Campbell, 2015; Mogensen, 2015). This hypothesis mirrors, to some extent, philosophical discussions of moral testimony. Some have argued that there is something distinctively problematic about deferring to others for moral action and judgment. In other words, they argue that there is something inherent to morality that requires individuals to give less weight to the advice or input of others (Aiken, 1962; Hills, 2009). This view makes intuitive sense: an individual should not have to defer to others to know that killing is wrong, and may be judged as not truly having a moral value without having arrived at that stance herself (Paul, forthcoming). Something similar may be driving people's judgments of decision-making. Deference may indicate inauthenticity, poor understanding of the situation on the decision-maker's behalf, or be perceived as reflecting a lack of careful consideration or care (Inbar, Cone, & Gilovich, 2010). These hypotheses could be tested in future studies by querying

people's judgments of what constitutes an authentic or careful decision-making procedure in a range of scenarios.

Finally, it is worth noting that the present set of studies did not actually survey people throughout their decision-making procedures. These results reveal an important aspect of folk epistemology, but more work is necessary to ensure its ecological validity. Future studies could query people who are in the process of making transformative choices, such as those deciding whether or not to have children, or those deciding whether or not to undergo sex reassignment surgery. This type of investigation would reveal whether people do actually privilege their own imaginations to the same extent that I previously observed. Further, it is important to understand whether this first-person privileging holds important post-choice consequences. For example, do people who privilege their own imaginations (as opposed to relying on testimonial or scientific evidence) have higher post-choice satisfaction, and lower post-choice regret? Do they feel more confident in their decision-making procedures?

Implications for the *real* limits of science

I have outlined so far the contributions of this research for how individuals perceive science and its limits, and how people privilege their own first-person knowledge when it comes to certain topics and types of decisions. It is important to emphasize, however, that this body of research reveals *intuitions* about the limits of science (and, more narrowly, intuitions among American adults) and does not actually speak directly to what science can or cannot explain, or the degree to which science actually can aid us in our decision-making. Nonetheless, this research does have important implications for how we think about the limited scope of science in our everyday lives.

At the outset, one important question worth posing is whether people's intuitions about the limits of science actually reflect the true state of the world. On the one hand, it could be that people's intuitions track some epistemic truth about the limits of science. This would mean, for example, that science is limited in its ability to explain aspects of the mind to which we feel that we have privileged first-person access, or that science has limited relevance when it comes to making transformative decisions. If these intuitions track epistemic truth, then science is truly limited in its explanatory scope, and, despite methodological advance, science will never be able to fully explain something like romantic love – especially what romantic love is like – because of its rich, first-person experiential quality.

On the other hand, if these intuitions about the limited scope of science are instead misguided, then they could prove to be serious barriers to scientific progress. People might think that other disciplines or other types of human endeavors – such as the arts, humanities, or religion – can better inform our understanding of romantic love and moral judgment, or better guide our decision-making in these domains. As Daniel Sarewitz wrote in his 2012 *Nature* commentary, “The Higgs boson, and its role in providing a rational explanation for the Universe, is only part of the story” (Sarewitz, 2012, p. 431).

Going one step further, however, if people falsely believe that a scientific perspective is not only insufficient but also misplaced or even harmful, then these intuitions have potentially large scale implications for the public acceptance of science and scientific progress. This might, at times, risk society at missing out on important truths, and potentially important scientific and medical advances. In recent years, for example, CRISPR and other gene editing technologies have gained much attention for their ability to transform medicine by eliminating previously incurable genetic diseases and disorders. In 2015, *Science* even referred to it as the

“breakthrough of the year” (Travis, 2015). Unsurprisingly, however, there has been great backlash to this sort of technology as well; some worry that gene editing could be used to create “designer babies,” or ultimately lead to eugenics programs. However, given people’s intuitive scope of science commitments – and the extent to which these commitments relate to our views on bioethics – it might be the case that these concerns reflect a more general discomfort with scientists pinpointing the genetic bases for many aspects of what it means to be human, thus coming worryingly close to overstepping its prescribed scope. Further understanding how these commitments influence judgment and decision-making would be important as we look toward how to best disseminate scientific results, and also how to most effectively teach science both in schools and to the general public.

Considering these two perspectives, which view is correct? Are people’s intuitions tracking some epistemic truth about the scope of science, or are they ultimately misguided? We may never know the answer to this question, but I am, at present, inclined to a middle position.

One theme that has emerged across these studies reflects people’s commitment that science cannot, or should not, explain or shed light on highly personal aspects of the mind or experience – the experiences to which we typically have introspective access, those that have a distinctive phenomenology, the ones over which we can exert conscious will, or those that are personally transformative. These intuitions are potentially onto something important. Perhaps there really are, in principle, certain aspects of experience that cannot be captured by scientific knowledge alone. Science benefits from its objective, third-person methodology, and this methodology – even if one day able to explain, for example, why Mary does or does not have the experiences that she does – supplies us with scientific knowledge, not personal experience.

However, it could also be the case that scientific explanations fall short of providing everything people want from them. It may not be that science actually fails to fully explain a phenomenon, but rather that science does not ground findings or explanations 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. As Virginia Woolf (1926) wrote: “The merest schoolgirl, when she falls in love, has Shakespeare or Keats to speak her mind for her; but let a sufferer try to describe a pain his head to a doctor and language at once runs dry” (p. 34).

References

- Aiken, H. D. (1962). Reason and conduct: new bearings in moral philosophy.
- Bear, A., & Knobe, J. (2016). What do people find incompatible with causal determinism?. *Cognitive Science*, 40(8), 2025-2049.
- Bear, A., & Knobe, J. (2017). Normality: Part descriptive, part prescriptive. *cognition*, 167, 25-37.
- Bloom, P. (2004). Natural-born dualists. *Edge: The Third Culture*.
- Bloom, P. (2009). *Descartes' baby: How the science of child development explains what makes us human*. Basic Books.
- Briggs, H. (2017). Normative theories of rational choice: Expected Utility, in E.N. Zalta (Ed.) *The Stanford Encyclopedia of Philosophy (Spring 2017 Edition)*, URL = <<https://plato.stanford.edu/archives/spr2017/entries/rationality-normative-utility/>>
- Budescu, D. V. (1993). Dominance analysis: A new approach to the problem of relative importance of predictors in multiple regression. *Psychological Bulletin*, 114(3), 542.
- Campbell, J. (2015). LA Paul's transformative experience. *Philosophy and Phenomenological Research*, 91(3), 787-793.
- Cappelen, H. (2012). *Philosophy without intuitions*. Oxford University Press.
- Chang, R. (2015). Transformative choices. *Res Philosophica*.
- Chudek, M., McNamara, R., Burch, S., Bloom, P., & Henrich, J. (2013). Developmental and cross-cultural evidence for intuitive dualism. *Psychological Science*.
- Denno, D. W. (2015). The myth of the double-edged sword: An empirical study of neuroscience evidence in criminal cases. *BCL Rev.*, 56, 493.
- Farah, M. J., & Murphy, N. (2009). Neuroscience and the soul. *Science*, 323(5918), 1168-1168.
- Forstmann, M., Burgmer, P., & Mussweiler, T. (2012). The mind is willing, but the flesh is weak: The effects of mind-body dualism on health behavior. *Psychological Science*, 23(10), 1239-1245.
- Goldenberg, J. L., Pyszczynski, T., Greenberg, J., Solomon, S., Kluck, B., & Cornwell, R. (2001). I am not an animal: mortality salience, disgust, and the denial of human creatureliness. *Journal of Experimental Psychology: General*, 130(3), 427.

- Gottlieb, S., & Lombrozo, T. (2017). Can science explain the human mind? Intuitive judgments about the limits of science. *Psychological science*, 29(1), 121-130.
- Greene, J. D. (2011). Social neuroscience and the soul's last stand. *Social neuroscience: Toward understanding the underpinnings of the social mind*, 263-273.
- Haidt, J., Bjorklund, F., & Murphy, S. (2000). Moral dumbfounding: When intuition finds no reason. *Unpublished manuscript, University of Virginia*.
- Haslam, N., Bain, P., Douge, L., Lee, M., & Bastian, B. (2005). More human than you: Attributing humanness to self and others. *Journal of Personality and Social Psychology*, 89(6), 937.
- Hatfield, E. (2006). The Golden Fleece Award: Love's Labours Almost Lost. *The Observer*, 19, 16-17.
- Hills, A. (2009). Moral testimony and moral epistemology. *Ethics*, 120(1), 94-127.
- Hitchcock, C., & Knobe, J. (2009). Cause and norm. *The Journal of Philosophy*, 106(11), 587-612.
- Horberg, E. J., Oveis, C., Keltner, D., & Cohen, A. B. (2009). Disgust and the moralization of purity. *Journal of personality and social psychology*, 97(6), 963.
- Inbar, Y., Cone, J., & Gilovich, T. (2010). People's intuitions about intuitive insight and intuitive choice. *Journal of personality and social psychology*, 99(2), 232.
- Inbar, Y., Pizarro, D. A., & Bloom, P. (2009). Conservatives are more easily disgusted than liberals. *Cognition and emotion*, 23(4), 714-725.
- Inbar, Y., Pizarro, D., Iyer, R., & Haidt, J. (2012). Disgust sensitivity, political conservatism, and voting. *Social Psychological and Personality Science*, 3(5), 537-544.
- Jackson, F. (1986). What Mary didn't know. *The Journal of Philosophy*, 83(5), 291-295.
- Kauppinen, A. (2007). The rise and fall of experimental philosophy. *Philosophical explorations*, 10(2), 95-118.
- Kumar, A. (2018). Disgust, stigma, and the politics of abortion. *Feminism & Psychology*, 0959353518765572.
- Mogensen, A. L. (2017). Moral testimony pessimism and the uncertain value of authenticity. *Philosophy and Phenomenological Research*, 95(2), 261-284.
- Myers-Schulz, B., & Schwitzgebel, E. (2013). Knowing that P without believing that P. *Noûs*, 47(2), 371-384.

- Nadelhoffer, T., Shepard, J., Nahmias, E., Sripada, C., & Ross, L. (2014). The free will inventory: Measuring beliefs about agency and responsibility. *Consciousness and Cognition, 25*, 27-41.
- Nagel, T. (1974). What is it like to be a bat?. *The philosophical review, 83*(4), 435-450.
- Nichols, S., & Knobe, J. (2007). Moral responsibility and determinism: The cognitive science of folk intuitions. *Nous, 41*(4), 663-685.
- Nimon, K., Oswald, F., & Roberts, J. K. (2013). yhat: Interpreting Regression Effects. R package version 2.0-0. Retrieved from <http://CRAN.R-project.org/package=yhat>.
- Paul, L. A. (2014). *Transformative experience*. Oxford University Press: Oxford.
- Paul, L. A. (forthcoming). Preference capture.
- Peer, E., Vosgerau, J., & Acquisti, A. (2014). Reputation as a sufficient condition for data quality on Amazon Mechanical Turk. *Behavior Research Methods, 46*(4), 1023-1031.
- Plunkett, D., Lombrozo, T., & Buchak, L. (2014). Because the brain agrees: The impact of neuroscientific explanations for belief. In P. Bello, M. Guarini, M. McShane, & B. Scassellati (Eds.), *Proceedings of the 36th Annual Conference of the Cognitive Science Society* (pp. 1180-1185). Austin, TX: Cognitive Science Society.
- Preston, J. L., Ritter, R. S., & Hepler, J. (2013). Neuroscience and the soul: competing explanations for the human experience. *Cognition, 127*(1), 31-37.
- Ranney, M. A., & Clark, D. (2016). Climate change conceptual change: Scientific information can transform attitudes. *Topics in Cognitive Science, 8*(1), 49-75.
- Richert, R. A., & Harris, P. L. (2008). Dualism revisited: Body vs. mind vs. soul. *Journal of Cognition and Culture, 8*(1), 99-115.
- Robinson, H. (2016). Dualism, in E.N. Zalta (Ed.) *The Stanford Encyclopedia of Philosophy (Winter 2016 Edition)*, URL = [<https://plato.stanford.edu/archives/win2016/entries/dualism/>](https://plato.stanford.edu/archives/win2016/entries/dualism/).
- Rozin, P., Haidt, J., & McCauley, C. R. (1999). Disgust: The body and soul emotion. *Handbook of cognition and emotion, 429-445*.
- Ryle, G. (1949). *The concept of mind*. University of Chicago Press: Chicago.
- Sarewitz, D. (2012). Sometimes science must give way to religion. *Nature News, 488*(7412), 431.

- Schwitzgebel, E. (2016). Introspection. In E.N. Zalta (Ed.), *The Stanford Encyclopedia of Philosophy (Winter 2016 Edition)*, URL = [<https://plato.stanford.edu/archives/win2016/entries/introspection/>](https://plato.stanford.edu/archives/win2016/entries/introspection/).
- Shariff, A. F., Greene, J. D., Karremans, J. C., Luguri, J. B., Clark, C. J., Schooler, J. W., Baumeister, R. F., & Vohs, K. D. (2014). Free will and punishment: A mechanistic view of human nature reduces retribution. *Psychological Science*, *25*, 1563-1570.
- Shenhav, A., Rand, D. G., & Greene, J. D. (2012). Divine intuition: Cognitive style influences belief in God. *Journal of Experimental Psychology: General*, *141*(3), 423.
- Simms, L. J., & Watson, D. (2007). The construct validation approach to personality scale construction. *Handbook of research methods in personality psychology*, 240-258.
- Slingerland, E., & Chudek, M. (2011). The prevalence of mind – body dualism in early China. *Cognitive science*, *35*(5), 997-1007.
- Stanovich, K. E. (1989). Implicit philosophies of mind: The dualism scale and its relation to religiosity and belief in extrasensory perception. *The Journal of Psychology*, *123*(1), 5-23.
- Strohming, N., & Nichols, S. (2014). The essential moral self. *Cognition*, *131*(1), 159-171.
- Strohming, N., & Nichols, S. (2015). Neurodegeneration and identity. *Psychological Science*, *26*, 1469-1479.
- Travis, J. (2015). Breakthrough of the Year: CRISPR makes the cut. *Science Magazine*.
- Uttich, K., & Lombrozo, T. (2010). Norms inform mental state ascriptions: A rational explanation for the side-effect effect. *Cognition*, *116*(1), 87-100.
- Wachbroit, R. (1994). Normality as a biological concept. *Philosophy of Science*, *61*, 579-591.
- Weisman, K., & Markman, E. M. (2017). Theory-based explanation as intervention. *Psychonomic Bulletin & Review*, 1-8.
- Wilson, T. D., & Gilbert, D. T. (2013). The impact bias is alive and well. *Journal of Personality and Social Psychology*, *105*(5), 740-748.
- Woolf, V. (1926). *On being ill*.