

Descartes' God is a Deceiver, and that's OK

Joseph Gottlieb & Saja Parvizian

Forthcoming in *Synthese*. Please cite published version.

Abstract

That Descartes' God is not a deceiver is amongst the canonical claims of early modern philosophy. The significance of this (purported) fact to the coherence of Descartes' system is likewise canonical, infused in how we teach and think about the *Meditations*. Though prevalent, both ends of this narrative are suspect. We argue that Descartes' color eliminativism, when coupled with his analysis of the cognitive structure of our sensory systems, entails that God is a deceiver. It's doubtful that Descartes recognized this, given his insistence that God is not a deceiver, and the role this plays in his system. But this is a concession Descartes can grant if we are careful about the kind of deception at play—a kind Descartes *does* recognize, albeit not without some ambiguity. On our story, Descartes' metaphysics and epistemology are not driven by concerns about deception *per se*, but by concerns about God's *benevolence*.

'You say that God cannot lie or deceive. Yet there are some schoolmen who say he can'.

- Mersenne

1 Introduction

We're often told that Descartes' God is not a deceiver. We're also told that God's non-deceptive nature secures the foundations of Descartes' philosophical system. That's the prevailing wisdom. We challenge it: in Descartes' system, God *is* a deceiver. And that's OK.

Central to our case is that Descartes must endorse:

SENSORY DECEPTION: Many sensations are deceptive such that they are apt to produce a certain kind of false sensory judgment—judgments about the apparent sensory qualities of objects—and their being so apt isn’t due to our violating various epistemic norms or the mere finitude of mind-body composites.

Our argument will focus on color sensations—sensory experiences *as of* colored bodies—in the face of Descartes’ *color eliminativism*.

Our thesis may induce dismissiveness. One may want to insist that no matter how the details are filled in, the argument will invariably be flawed, or SENSORY DECEPTION is itself false. Never mind that Descartes often *flatly denies* that God is a deceiver. God cannot be a deceiver, the thought goes, because God is omniperfect. And with respect to our sensations, Descartes claims that God was benevolent in giving them the function of preserving our health (Sixth Meditation, AT VII: 83/CSM II: 57), and giving us the best of all possible sensations to that end (Sixth Meditation, AT VII: 87-8/CSM II: 60).¹

Now on the one hand, any ties between being omniperfect and being a non-deceiver, or between a type of sensation being optimally suited to preserve our health and its being non-deceptive, is itself part of what needs to be worked out. On the other hand, if God were the cause of a *malicious* or *harmful* form of deception, that would obviously be deeply problematic, so much so that this dismissive stance would be warranted.

These observations lead to two claims. First: if God is a deceiver, he better not be a malicious deceiver. Second: it’s question-begging to assume that any form of deception is incompatible with God’s omniperfection. Not everything that is deceptive is maliciously deceptive. And color sensations, we’ll claim, are such a case. They are non-intellectually, *non-maliciously* deceptive; non-intellectual because they are only partially grounded in the intellect (a body is needed too—see *Principles* I.32, AT VIIIA: 17/CSM I: 204), and non-malicious because their being deceptive does not cause harm.²

Our precisified thesis, then, is that God is a non-malicious deceiver, and that if Descartes were to admit *this*, it would be OK.³ After setting out our argument, we’ll look at four objections—three concerning whether we may be responsible for our false color judgments, not God, and one concerning the unique kind of color judgments at play—arguing that these

¹We employ the following abbreviations: ‘AT’: *Oeuvres de Descartes* (cited by volume and page), Adam and Tannery (1996); ‘CSM’: *The Philosophical Writings of Descartes* (cited by volume and page), Cottingham et al. (1985); ‘CSMK’: *The Philosophical Writings of Descartes* (cited by page), Cottingham et al. (1991).

²As we will see, there might be other ways in which deception conflicts with God’s omniperfection—e.g. via potential conflicts with God’s omnipotence. However, our focus will be on the relationship between malice, harm, and omnibenevolence. Although see fn. 33.

³Patterson (2016) also endorses SENSORY DECEPTION and that SENSORY DECEPTION poses a problem for God’s non-deceptive nature. However, she claims that Descartes does not have a solution to this problem. We argue otherwise.

all fail. Dialectically, the key question for us is not how Descartes' views about God may or may not constrain his views about our sensory systems. The question here is what Descartes *in fact says* about our sensory systems; and what he says entails that God is a deceiver, albeit a non-malicious one.

In view of this, one might just say that, even if our argument is sound, Descartes was simply inconsistent, not realizing what his views about our sensory system committed him to. That's certainly plausible, and not an option we attempt to rule out here. Instead, we attempt to explore whether Descartes *could* have allowed for divine deception, irrespective of whether he actually did, and what this might look like. Without making any claims about what Descartes definitively meant, we'll suggest that Descartes has the resources to allow for our 'omni-compatible' sense of divine deception, while maintaining the integrity of his core arguments in the *Meditations*. As Mersenne tells Descartes, various medieval philosophers allowed for divine deception (Perler 2010); we suggest that Descartes can too.

2 The Argument

Descartes only offers examples of deception (see, e.g., First Meditation, AT VII: 18/CSM II: 12, *Discourse I*, AT VI: 9/CSM I: 115, *Optics* AT VI: 145/CSM I: 174). Even so, given his examples, Descartes would endorse:

DECEPTION: *X* deceives a subject *S* if (i) *S* judges that *p*; (ii) it's false that *p*; (iii) *X* judges that *p* is false; (iv) the evidence *E* on the basis of which *S* judges that *p* is brought about by *X* for the purpose of making *S* judge that *p*.

DECEPTION captures a form of deception by commission.^{4,5} However, as we will see, whether God is a deceiver under DECEPTION depends on how 'judge' ('judgment') is understood.

Our argument to follow goes like this:

- D1 Bodies are not colored.
- D2 In virtue of bodies looking colored—i.e. in virtue of the fact that certain phenomenological evidence *E* obtains—*S* judges that bodies are colored.
- D3 So, in virtue of bodies looking colored—i.e. in virtue of the fact that certain phenomenological evidence *E* obtains—*S* *falsely* judges that bodies are colored.
[D1-D2]

⁴See Mahon (2015).

⁵Dropping 'for the purpose of' from clause (iv) makes room for *non-intentional* forms of deception (Chisholm & Feehan 1977). Everything to follow will apply on non-intentional deception too, since (iv) entails its weaker non-intentional, purely causal cousin.

- D4 God brought about E for the purpose of making S falsely judge that bodies are colored.
- D5 God knows that bodies are not colored.
- D6 If D3, D4, and D5, God is a deceiver [given DECEPTION].
- ∴ God is a deceiver.

D5 is trivial, given color eliminativism (D1) and God’s omniscience. Since D3 is a bridge conclusion, what remains to be discussed are D1, D2, D4, and D6.

2.1 Regarding D1

D1 says that bodies are not colored. If Descartes endorsed D1, he was a color eliminativist.⁶

Defenses of Descartes’ endorsement of D1 can proceed in different ways. One is to move from the ground up, with a direct, independent argument appealing to primary sources. Our ambitions have limits, so readers must look elsewhere for such an undertaking. To that extent, D1 will be taken for granted. That’s hardly an assumption worth blushing at though, with eliminativism arguably being the most common interpretation not just of Descartes’ views on color, but across early modern philosophy more generally, including Galileo, Hobbes, Boyle, Malebranche, and Locke.⁷ The rational is familiar, stemming from *mechanism* about bodies, (e.g., AT VII 440/CSM II: 297, *Principles* II.4, AT VIIIA: 42/CSM I: 224), so we won’t recount it here.

What’s especially curious for present purposes, however, is how both proponents and opponents of the argument from mechanism—or just proponents and opponents of D1 more generally—often display a relaxed attitude when it came to color terminology, allowing that color terms like ‘red’ and ‘green’ are equivocal or ambiguous.⁸ This yields a second way of defending D1, albeit one that is indirect. Boyle once noted that the “famous controversie” between color eliminativism and color realism are “in great part a nominal dispute” (1664: 75).⁹ So perhaps those in the secondary literature like Cottingham (1989) and Atherton

⁶D1 could have been formulated differently, as, e.g. “Bodies do not have tastes”. In this way, our argument requires only eliminativism about *a* sensory quality, not color eliminativism per se. The focus on color has theoretical and practical motivation; theoretical because color experiences deliver arguably the least ambiguous phenomenological predictions (making assessments of non-veridicality more tractable), and practical because Descartes had more to say about color than other sensory qualities and experiences thereof.

⁷For a sampling of eliminativist readings, see: Buroker (1991), Chamberlain (2019a), Cook (1996), Downing (2011), Garber (1992: 75), Keating (2004), Maund (2011: 382), and Rozemond (1999: 452-3). The Stanford Encyclopedia of Philosophy entry on color also describes Descartes as an eliminativist.

⁸Schmaltz (1995: 391), for example, mentions how Descartes and Malebranche held that colors terms could apply not just to sensations or visual experiences, but also to their bodily causes. For further discussion, see Chamberlain (2019a).

⁹See Chamberlain (2019a) for this reference.

(2004) who appear to be arguing that Descartes rejected D1, are actually advancing a reading that is consistent with D1. Descartes may very well have agreed that grass is green in some sense of ‘green’, but nothing cuts against the idea that grass isn’t green in the sense of ‘green’ that actually matters for whether Descartes’ God was a deceiver. To see how this might work requires a closer look at eliminativism itself, along with its motivations in the contemporary literature.¹⁰

What unites all forms of eliminativism is this: external bodies—Granny Smith apples, grass, and national flags—do not instantiate colors. From here, there are options.

Some versions of eliminativism say that colors are instantiated in the mind or the visual system, as on *literal projectivism* (Boghossian & Velleman 1989; Averill 2005). Others say that colors are wholly uninstantiated, as on *figurative projectivism* (e.g., Averill 1992; Wright 2003). Others still deny that colors exist in any way, even as uninstantiated universals (Hardin 1988), as on *strong eliminativism*. Both forms of projectivism are weak forms of eliminativism insofar as they are better understood as *relocating* colors (to the head, or to the realm of abstracta) rather than denying their existence outright, but not much is in a name.

Further, while all eliminativists agree that bodies *look* or *appear* colored, an eliminativist can say what they wish about these appearances—e.g. that they are not perceptual, but doxastic. Interestingly, Nolan (2011) says that Descartes endorsed something like this view. According to Nolan, Descartes is a color nominalist. On Nolan’s reading, Descartes often prefixes his purportedly positive metaphysical claims about color with ‘what we call’ to signify that he is not making any positive metaphysical claims about the nature of color, but merely semantic claims about how color terms are normally used. And while *color appearances* are real, they are not a sensory or perceptual affair (they are not “taught by nature”), but are rather due to voluntary, habitual, false judgments (Ibid. 84, 98). These judgments have at least two parts: that bodies are colored, and that colors are qualities (Ibid. 107). So these habitual judgments are actually false twice over. It’s not just that bodies are not colored. According to Nolan, colors are not even qualities.¹¹ And that implies that colors *do not exist*, since if they did, they would be qualities (i.e. properties). Saying that bodies are not colored suffices all on its own for eliminativism, but this latter claim makes Descartes

¹⁰We thank an anonymous referee for encouraging us to say more here.

¹¹Nolan (2011: 107): “these false judgments include not only the Scholastic view that ‘colors’ are qualities in objects, but also the more fundamental position that colors are ‘qualities’ at all”. That colors are not qualities is supposedly buttressed by Descartes’ warning to Burman that we err if we attempt to ‘reify’ p-colors in any way (AT V: 152/CSMK: 337). Note, however, that *reification* is simply *to make concrete*. So, we could in fact heed Descartes’ warning while still saying colors exist—they would just exist as uninstantiated properties, as on figurative projectivism. Of course, if Descartes were a figurative projectivist, he would have to reject an Aristotelian view of universals, where universals are *concreta-dependent*. While treating Descartes as a Platonist about universals has precedent in the secondary literature (e.g., Schmaltz 1991), this is naturally not without controversy (e.g., Nolan 1998).

a strong eliminativist, albeit one who grounded the appearance of bodies as colored in habitual judgments. As we will see in our discussion of D4, we doubt that grass looks green because of habitual false judgments. Critically, this matters for whether Descartes' God is a deceiver—an issue that Nolan (Ibid. 98) explicitly recognizes. For now though, the main point is that all D1 requires is that Descartes was an eliminativist of some stripe.

We suspect, however, that any confusion over D1 arises not from eliding the distinction between weak and strong eliminativism, or the different ways in which the eliminativist might analyze color appearances. The crux concerns the property the eliminativist is trying to 'eliminate' or 'relocate;', and *what such elimination and relocation may be consistent with.*

The eliminativist is interested in our ordinary concept of color (Maund 2011). Our ordinary concept of color concerns what Mackie (1976) calls 'colors-as-we-see-them', or what we will call *phenomenal colors* ('p-colors'). P-color concepts are recognitional concepts. Employing a recognitional concept for F involves telling that x is F by the way x looks to you when one has an experience of x. This means that the application of color names will be fixed by the way objects appear, and the properties color terms ('green') name is just a property that objects appear to have (Maund 2011: 365, cf. Byrne & Hilbert 2003).¹² So, what fixes the subject matter for the sort of color that is to be eliminated or relocated is whatever properties our ordinary color concepts and color naming practices (putatively) pick out (Ibid. 366). These properties are p-colors: the often conspicuous, qualitative features bodies appear to have. And according to the eliminativist, our experiences of bodies as p-colored are not just occasionally non-veridical, but always and systematically so; they misrepresent bodies as p-colored, and they do so in a consistent and stable way.^{13,14}

Now defenses of eliminativism in the contemporary literature often proceed by consulting common sense beliefs about p-colors (e.g., Johnston 1992; Maund 2011), before arguing that scientific investigation reveals that bodies do not instantiate anything that meets this common-sense profile. Yet this is consistent with bodies instantiating other properties that

¹²Nolan tells us that Descartes concurs: "Descartes thinks that we ordinarily pick out the properties of bodies that cause or sensations on the basis of the sensations themselves" (2001:106). In other words, by Nolan's lights, Descartes would agree that our p-color concepts are recognitional concepts.

¹³Going forward, we shorten the expression 'phenomenal color' to 'p-color' and refer to all determinate phenomenal colors by prefixing them with a 'p'; p-red, p-green, and so. There will be cases where using 'color' will be unavoidable, either because using 'p-color' begs the question, or because 'color' is used in the source material. See Chamberlain (2019a) for a similar terminological regimentation; where he uses 'sensuous color', we use 'p-color'.

¹⁴In saying that color experiences misrepresent bodies, we are assuming *representationalism* (Simmons 1999, De Rosa 2007; Hatfield 2013). Indeed, we think that, for Descartes, *all* experiences represent, at least in the broad sense so as to include non-indicative contents (Gottlieb & Parvizian 2018). But nothing substantive here turns on this, even for color experience specifically; *qualia realists* (MacKenzie 1990) can grant color eliminativism. They just wouldn't cash out bodies looking p-colored in terms of their being (mis)-represented as p-colored.

are related to p-colors. Byrne and Hilbert, who endorse a version of color physicalism, make the point this way:

At the start of enquiry, one would want to make a distinction between salt and sodium chloride...even though it may turn out that salt is sodium chloride...It may similarly turn out with phenomenal color and (a kind of) physical color. Although care must be taken to make this distinction at the outset, perhaps phenomenal and physical color are one and the same (2003: 6).

The idea is that, after thorough investigation, ‘physical color’—for Byrne and Hilbert, an anthropocentrically demarcated disjunction of spectral reflectances—might turn out to be p-color. The eliminativist says it won’t. However, the eliminativist doesn’t thereby deny that bodies reflect light in certain ways.

The same point applies to *dispositionalist* views (e.g., Cohen 2009). Like the physicalist, the dispositionalist rejects eliminativism and identifies p-colors with dispositions to affect certain perceivers in certain ways in certain contexts, dispositions which bodies possess. Yet the eliminativist does not, or at least need not, respond to this by denying that bodies have these subject-involving dispositions (Averill & Hazlett 2010: 761). The eliminativist will (or can) grant that bodies have ‘physical colors’ and will (or can) grant that bodies have ‘dispositional colors’. What she denies is that these are p-colors.

These points might seem simple enough, but they reveal something crucial about the secondary literature on Descartes’ views about colors. A theory of color is in the first instance a theory of p-colors. It’s not a theory of physical colors, except insofar as one may be inclined to identify p-colors with physical colors, nor is it a theory of dispositional colors, except insofar as one may be inclined to identify p-colors with dispositional colors. Now consider this passage:

[W]e have every reason to conclude that the properties in external objects to which we apply the terms light, color, smell, taste, sound, heat and cold, as well as other tactile qualities...are, so far as we can see, simply various disposition in those objects which make them able to set up various kinds of motion in our nerves. (*Principles* IV.199, AT VIII: 322/CSM I: 285)

This passage has led some—most notably John Cottingham (1989)—to argue that Descartes was a dispositionalist. This would lead one to think that, on Cottingham’s reading, Descartes identifies p-colors with dispositional colors. But this is hardly clear. This passage does, of course, suggest that Descartes identifies *some* kind of color property—“are...simply”—with such dispositions. Yet it’s a further claim that these dispositional properties are p-colors. And Cottingham seems to in fact *deny* this further claim:

What is denied is the inherence of redness qua redness—redness construed as a certain sort of sui generis quality supposed to inhere in objects in a way that exactly matches our sensory awareness of it. (1989: 238)

What Cottingham is telling us is that there is nothing in bodies that is like the property of which we have awareness. This, presumably, is what “exactly match[ing]” comes to: there is no property in bodies that has all the essential features of colors-as-we-see-them.¹⁵ And this property of which we have sensory awareness is a property ‘supposed to inhere in objects’—it’s a property that bodies appear to have. These are p-colors. So, yes, there is a property F, a dispositional property, that bodies have, but there is this other property F*, of which we have sensory awareness, that does not match F. But that is just to be an eliminativist with respect to F*, i.e. p-colors.

Something similar is afoot in Margaret Atherton’s (2004) reading. She contends that Descartes uses ‘color’ to pick out different properties in different contexts. For example, in the *Optics*, Descartes writes:

Colors are nothing other than the various ways in which the bodies receive light and reflect it against our eyes. (AT VI: 85/CSM I: 153)

This passage suggests one kind of physicalism: colors are spectral reflectances. Elsewhere, Descartes suggests another kind of physicalism: colors are tendencies for bodies to rotate in certain ways (*Meteorology*, Eighth Discourse, AT VI: 333). But Atherton seems to admit that Descartes was not a physicalist *where it counts*, viz. for p-colors:

What makes physical states have the character they have is not the same as what makes psychological states have the character they have. The green you see is not the same as and is not very much like the green in the world. (Ibid. 33)

The ‘green’ in “the green you see” is p-green. ‘See’ is a success-term. So if we see p-green, p-green exists. But bodies only instantiate physical green. So p-green isn’t physical green. Hence, since bodies still appear p-green, and only appear p-green, it follows that Descartes was an eliminativist. That bodies have physical colors is neither here nor there.

We don’t deny that there are tricky questions about just when a proposed reduction (of, say, p-colors) is a genuine reduction or is really an elimination. Our claim here isn’t

¹⁵A claim in the background here is what Johnston (1992) calls *Revelation*—roughly, the idea that the intrinsic nature of p-colors is fully revealed in our experience of them. Notice that eliminativism as such isn’t committed to p-colors having no hidden nature. The literal projectivist will say that p-colors are instantiated in our visual system, but we are not aware of p-colors as being our heads in this way. We are aware of them *as* being located on the surface of external bodies. In this way, the “exactly matching” can just be read as saying that *of* the features we are aware of p-colors as having, there is no property in bodies that matches a property of this description.

that it's *obvious* that physical and dispositional colors are not p-colors. Still, as Gow (2014) points out, it's significant that both physicalist and dispositionalist views do indeed seem to give us accounts that move us far from our common-sense conception of color, and that the eliminativist can agree that bodies possess the kinds of properties picked out by physicalists and dispositionalists.¹⁶ This leaves us to wonder, as Gow (Ibid.) also suggests, whether the dispute between eliminativism and these forms of realism is just terminological: the disagreement is over which kinds of properties get to be called 'colors'.

In short, it is *p-colors* that appear to us in visual experience as properties of bodies, so it is the instantiation of *those* properties in bodies that matters for whether God is a deceiver. We grant D1 is not wholly innocuous, and interpretive questions are not monopolized by the primary texts. But we hope it is plausible enough to make our thesis worthwhile.

2.2 Regarding D2

D2 concerns *sensory* judgments. Broadly, a 'sensory judgment' refers to a judgment about some sensory state of affairs. A sensory state of affairs can concern just sensible qualities (for example, the p-redness of a rose), sensations themselves, or both. Not all sensory judgments need to be immediately caused by a sensation. I could, for instance, judge that roses are red based on my memory of seeing a rose. However, in the *Sixth Replies* Descartes recognizes a class of sensory judgments—one of the three grades of 'sensory response'—that by their nature are immediately caused by sensations. The first grade of sensory response is purely physiological and consists of the "immediate stimulation of the bodily organs by external objects" and terminates in a brain impression (AT VII: 436/CSM II: 294). The second-grade of sensory response occurs in the mind, and comprises "all the immediate effects produced in the mind as a result of its being united with a bodily organ which is affected in this way" (Ibid.). These are 'sensations' in our parlance. The third-grade of sensory response involves cognitive processes, and "includes all the judgements [*judicia*] about things outside us [*de rebus extra nos*] which we have been accustomed [*consuevimus*] to make from our earliest years—judgements which are occasioned by the movements of these bodily organs" (AT VII: 437/CSM II: 294-5).

The operative word here is 'occasioned'. Let *third-grade judgments* be judgments about sensory states of affairs—specifically, sensible qualities—that are immediately occasioned by events at the first and second-grade of sensory response. We think third-grade judgments

¹⁶For a discussion of how these views depart from our common-sense conception of color, see Gow (2014), who offers that the only realist views of p-color are primitivist views, since they at least agree that bodies have the properties that we ordinary take to be colors. However, Chamberlain (2019a: 299) argues that Cavendish was a color physicalist, who held that "grass is green in precisely the way it visually appears to be". It's of course possible that Descartes held such a view, but Atherton seems to be denying it.

have two unique features. First, they in some way structure how things look phenomenologically. Second, they are irrevisable. Following Simmons (2003), we will refer to third-grade judgments as *constructive judgments* (‘c-judgments’) to emphasize this first feature. So defined, not all sensory judgments are c-judgments. Some sensory judgments might neither structure phenomenology nor be irrevisable. These are *reflective judgments* (‘r-judgments’). The existence of r-judgments should be uncontroversial, but c-judgments also have gained broad favor in the secondary literature. Simmons (Ibid.), Hatfield (1986), and Patterson (2016) all contend that Descartes countenanced c-judgments.¹⁷

Now D2 says that S judges that bodies are p-colored in virtue of bodies looking p-colored (i.e. E). While we do make r-judgments to this effect, the sense of judgment relevant to D2, and our case that God is a deceiver, concerns c-judgments. D3 (that S *falsely* c-judges that bodies are p-colored) follows from D1 and D2. We will thus take up three claims: first, that Descartes countenanced c-judgments; second, that c-judgments about p-color are false; and third, following Patterson (*pace* Simmons) c-judgments are issued by the will. A corollary of the third claim concerns the relationship between c-judgments and beliefs.

Claim One. Let’s start by looking again at Descartes’ three grades of sensory response, focusing on p-color. At the first grade, there is some property in bodies—perhaps a dispositional property or the categorical basis of this dispositional property—that leads to a brain impression that, in turn, causes a p-color sensation. The sensation itself is the second-grade of sensory response. It’s in virtue of this second-grade response that bodies *look* colored. This is because sensations *represent* bodies as p-colored. That sensations represent follows from Descartes’ general claim that all mental states have objective reality, that they are “of things [*rerum*]” (Third Meditation, AT VII: 44/CSM II: 30), and his more specific claim that sensations inform [*significandum*] (Sixth Meditation, AT VII: 83/CSM II: 57). Given eliminativism, this implies that bodies are misrepresented as colored. So, again, when it comes to D2, and its talk of ‘phenomenological evidence’, this concerns the second-grade of sensory response; it’s the misrepresentation at the second-grade that constitutes the phenomenological evidence (E) in virtue of which S is deceived, if in fact S is deceived.

At the third-grade, the mind makes a host of judgments about what occurs at the second-grade. The question is whether these judgments play their own role in structuring sensory phenomenology and are irrevisable—that is, whether there are c-judgments so defined. Here, two points are especially germane:

- (1) Third-grade judgments partially structure sensory phenomenology.
- (2) If third-grade judgments partially structure sensory phenomenology, then given

¹⁷When unqualified, ‘judgment’ can refer to a c-judgment or r-judgment.

Descartes' views about the status and functional role of our God-given sensory system, third-grade judgments are also irrevocable.

If (1) and (2) hold, then third-grade judgments are c-judgments so defined.

Extensive arguments for (1) have been offered by Simmons (2003), Hatfield (1986), and Patterson (2016), so we'll confine ourselves to some clarificatory remarks. First, most of the examples of c-judgments adduced by Simmons, Hatfield, and Patterson pertain to size, shape, and distance. However, Simmons is explicit that there are c-judgments about p-color as well, writing that “[a]s a result of the intellect’s constructive judgments, bodies are represented. . . as having constant [p-]colors and shapes that are viewed under changing conditions” (2003: 575). Second, it would be odd if Descartes did not countenance c-judgments about p-color if he countenanced c-judgments about size, shape, and distance. Presumably, there is some sort of feed-forward mechanism from what’s represented at the second-grade level of sensation to what’s represented at the third-grade level of c-judgments, even if the latter representations have a different format (e.g. discursive and propositional) than the former (e.g. iconic and non-propositional).¹⁸ And since size, shape, distance, and p-color are all presumably at the second-grade, it would be surprising if the former three get taken up at the third-grade but not the latter. Indeed, on the assumption that Descartes is talking about such judgments in the aforementioned passage of the *Sixth Replies*, he seems to take such an inclusive stance, saying that these judgments include “*all* judgements about things outside us....judgements which are occasioned by the movements of these bodily organs” (emphasis added). And so, since p-color sensations concern things outside us (bodies) and are occasioned by movements of the bodily organs at the first grade of sensory response, there will therefore be judgments that structure the phenomenology of p-color experience at the third grade of sensory response, if there are such sensory judgments at all. This is consistent with the *Sixth Replies* discussion of the three grades of sensory response, where Descartes writes that “as a result of being affected by this sensation of colour, I judge that a stick, located outside me, is coloured” (AT VII: 437/CSM II: 295).

Second, we make no claims about how exactly these judgments structure phenomenology. In the case of judgments about shape, Simmons tells us that:

When I look at a bagel from an oblique angle, I judge that it is round (or perhaps toric) despite the elliptical patch of brown it produces in my visual field at the second-grade of sensory perception; the result is a sensory experience in which the bagel looks round (or toric). (2003: 554)

¹⁸According to Descartes, sensory perception falls under the faculty of the intellect, which supplies “ideas which are subjects for possible judgement,” where the judgments are issued by the will (Fourth Meditation, AT VII: 56/CSM II: 39). In the context of the grades of sensory response, we take “subjects for possible judgment” to roughly mean: what is represented by second-grade sensations.

With p-color sensations, Simmons says that these judgments account for p-color constancy, where bodies are represented as having a particular p-color across distinct illumination conditions (2003: 573). As Simmons writes: “I may therefore judge, at the third-grade of sensory perception, that those leaves are green *with the result that* the leaves look like green leaves in changing light” (Ibid.; emphasis added). Whether or not this is the correct story, our claim is just that there are judgments that structure the phenomenology of sensory experience, and this includes experiences of p-color.

Regarding (2), to say that a third-grade judgment is irrevivable is to say that one cannot cease the formation of a third-grade judgment, or alter its content-type, once a second-grade sensation is had. To say that they are revisable is to say otherwise. The case for (2) is broadly theoretical. The case, inspired by Patterson (2016), goes like this. Descartes claims that our sensory faculties—including *all* grades of sensory response—have a biological function of preserving the health of mind-body composite. As such, it’s not just the sensations at the second-grade that play this role, but third-grade judgments as well. And critically, while both the second and third-grades work in concert to preserve our health, the *way* they do this will be different. *Whatever* phenomenology-determining role third-grade judgments play—and we know that they play some role, given point (1)—it’s *that very role* that helps preserve our health. For example, if Simmons is right, the second-grade makes bodies look p-colored *simpliciter* (something that is good for us), and the third-grade makes bodies look to have the *same* color across different illumination conditions (also good for us). More generally, if God gave us the best of all possible sensory systems, then the way third-grade judgments structure phenomenology must be perfectly suited towards our preservation. In other words, the third-grade judgments we *in fact* issue have to be the best of all possible such judgments *qua* their bio-functional role. If God gave us the best of all possible sensory systems, and such judgments are part of that system and so figure into the preservation of our well-being, then any attempt to revise them would, given our fallibility and finitude, *ipso facto* make them worse. Thus from a design standpoint, it makes more sense for these judgments to be irrevivable. It’s no wonder then that they are considered automatic, and caused directly by sensory states without any inference.¹⁹

¹⁹Simmons take on irrevivability is different. For Simmons, c-judgments are an essential constituent of sensory experience, and since sensory experiences are themselves irrevivable, she concludes that third-grade judgments are irrevivable (2003: 578). We are not inclined towards this line, since we think there is a good sense in which the second-grade constitutes a sensory experience all on its own. Simmons reserves ‘sensory experience’ for the joint product of the second-grade sensation and a c-judgment. We reject this usage. This is not to deny that c-judgments have a unique role, as Simmons contends. Since on our view the second-grade sensation is sufficient for a body to look colored *simpliciter*, this strikes us as good of a candidate as any for being a sensory experience. What we insist on is that there is no *further* thing that results when second-grade sensations and c-judgments co-occur.

R-judgments, by contrast, are revisable (Simmons 2004: 567-8; cf. Patterson 2016: 99). They are formed through inference. The meditator judges that “the heat in a body is something exactly resembling the idea of heat which is in me” (AT VII: 82/CSM II: 56). This judgment is revised—I can cease to have it altogether—once the meditator learns that heat isn’t a property of extension. R-judgments don’t structure sensory phenomenology, and so are not part of our sensory system in the strict sense.

Claim Two. Are c-judgments false? Simmons seems to think so, saying that, when it comes to judgments about size, distance, shape, *and color* that “in none of these cases does the intellect produce a properly intellectual representation of the corporeal world, or even a reliably accurate representation of it” (2003: 575).²⁰ But our case doesn’t depend on such a strong claim. All that matters is that c-judgments about p-color are false. And they *must* be, if eliminativism is true. When I c-judge that the stick, located outside me, is p-colored, I am c-judging falsely, because nothing (including the stick) is p-colored. So, our second claim is indeed conditional on Descartes’ being an eliminativist. We trust that there are far worse things to stake one’s case on.

One might wonder though whether the content of my c-judgment is more like <the stick looks p-brown> as opposed to <the stick is p-brown>. After all, the former c-judgement is true. However, that’s not what Descartes says we c-judge. He says that we c-judge that “a stick, located outside me, is coloured”. And this is arguably what Descartes *should* say anyway.

To see why, suppose that S has a sensory experience of the stick as p-brown. Now assume, as we are, that c-judgments help fix experiential phenomenology. If, as the alternative suggests, it was the case that:

- (3) S c-judges that <the stick looks p-brown>.

then, in describing the phenomenology of her experience, S would say that:

- (4) It looks as if the stick looks p-brown.

But this is manifestly *not* the correct description of the phenomenology. The correct description is just:

- (5) It looks as if the stick is p-brown.

²⁰Patterson (2016: 101) also seems to think so, saying that “[f]alse estimates of distance and erroneous constructive judgements about shape and size do not happen occasionally, when conditions are unusual or when the body is damaged or diseased. They occur routinely, as a consequence of the ordinary functioning of the visual system.”

We get (4) instead of (5) since it's the *content* of the c-judgment—here, <the stick looks p-brown>—that does the phenomenological work, and the content in (3) includes the property of *looking*. Yet it's doubtful that (4) describes a real sensory-level phenomenological fact. The stick (visually) looks p-brown. This is just because I am having a sensory experience of the stick as p-brown, and that's *one and the same thing as* the stick's (visually) looking p-brown. In this way, the 'It looks' in (4) and (5) function as experiential operators. Yet the property of *looking* is not *itself* phenomenologically salient, contra (4); the (visually) looking is just the experiencing, and what's phenomenologically salient in the experiencing is the apparent *p-brownness* of the stick, as in (5). So if c-judgments have a role in fixing the sensory-level phenomenological facts, they won't look like (3), and so they won't be true on that count.

Claim Three. For Simmons, r-judgments are not just different from c-judgments because the former, but not the latter, are irrevocable. They are also different because the latter are not beliefs, even if they often lead to them (Simmons 2003: 567). A judgment is a belief only if it is issued by the will. Since Simmons denies that c-judgments are issued by the will, she claims they are not beliefs. Rather, they are a product of the intellect.

As to the *source* of c-judgments, Simmons realizes that her take conflicts with Descartes' standard theory of judgment in the Fourth Meditation, on which judgments are issued by the will (2003: 565-6). Simmons supports this alternative reading by drawing from Descartes' claim in the *Sixth Replies* that "although such reasoning is commonly assigned to the senses (which is why I have here referred it to the third-grade of sensory response), it's clear that it depends solely on the intellect [*a solo intellectu pendere*]" (*Sixth Replies*, AT VII: 437-8/CSM II: 295; cf. *Optics* (AT VI: 137-8/CSM I: 170).

Yet as Patterson points out, Descartes' use of 'intellect' here is misleading:

I read Descartes's talk of the intellect alone as designed to emphasise that the senses are not involved, rather than to exclude any role for the will. This...gains support from Descartes associat[ing] judgments in the full-blooded sense with the 'intellect alone' at the end of the Second Meditation, where his point is also to contrast judgment with the senses and imagination...Descartes uses 'intellect' as an umbrella term to cover intellect and will, the faculties of pure mind, when a contrast is being made with the faculties of the embodied mind. (2016, fn. 48)

Thus we'll assume r-judgments and c-judgments are issued by the will.²¹

²¹Descartes' claim that sensations are wholly passive is consistent with c-judgments being issued by the will. This *prima facie* inconsistency is in part due to Descartes' slippery usage of 'sensation'. As evidenced by his clarifications in the *Sixth Replies*, he sometimes uses 'sensation' to refer to the composite sensory experience that includes the third and second-grade of sensory response, and sometimes he uses 'sensation' to refer to the second-grade alone (our usage). Resolving the tension is straightforward: sensations *at the*

Was Simmons right to deny that c-judgments are beliefs? There are several issues here. First, while Descartes held that being issued by the will is necessary condition on belief, it's unclear if it is sufficient. So our disagreement with Simmons about the source of c-judgments doesn't settle things. Second, while Descartes sometimes seems to use 'belief' and 'judgment' interchangeably (e.g., AT VII: 58-9/CSM II: 41; AT VII: 69/CSM II: 48), this doesn't settle things either, since *c*-judgments are rather curious as a mental kind. We have at least a decent handle on beliefs. They have propositional content. They are assertoric in force. They figure in inferences. They are linguistically expressible in suitable creatures, and they characteristically interact with desires, emotions, action, and memory (Quilty-Dunn 2015: 559). Yet it's unclear if *c*-judgments check all of these boxes. For example, Simmons (2003: 554 fn. 13) flirts with the idea that *c*-judgments are "confusedly" presented to consciousness (*Principles* I.46, AT VIII-A 22/CSM I: 208), which might make them linguistically inexpressible. On the other hand, as Quilty-Dunn notes (2015: 559), since beliefs are functional states, and so specified in terms of roles they are *disposed* to fulfill, that *c*-judgments may not *actually* fulfill these roles is otiose, as "dispositions can be muted" (Ibid.)

At the end of the day, we hope to be neutral on this issue, as we do not think this choice-point matters for the main question at hand, viz. whether God is a deceiver.²² We will return to this point in our discussion of D6. For now, we'll just note that at the very least, *c*-judgments are 'belief-like' insofar as they are assertoric in force and propositionally truth-apt. Judging that *p* is "the act of occurrently putting *p* forward in one's mind as true" (Cassam 2010: 83). So one can thus *c*-judge falsely, or put *p* forward in one's mind as true when it isn't. As will become clear, the existence of *c*-judgments is what renders Descartes' God a deceiver. What is *deceptive* are second-grade p-color sensations, but S is only *deceived* once she makes false *c*-judgments.

2.3 Regarding D4

D2 claimed that S makes *c*-judgments to the effect that bodies are p-colored due to certain phenomenological evidence E—i.e. the phenomenological fact that bodies look p-colored at the second-grade. D4 goes further. Adjusted for subsequent clarifications, it says:

second-grade are wholly passive. One might also worry that if we cannot help but issue *c*-judgments, then the will cannot be wholly free. Yet it's fairly clear that Descartes recognizes *conditional* constraints on the will. For instance, Ragland (2006) contends that Descartes endorsed this claim: given certain conditions, *if* S has a clear and distinct perception that F, S cannot refrain from judging that F (Fourth Meditation, AT VII: 58-9/CSM II: 41). Yet this is structurally analogous to our claim about *c*-judgments: we only claim that *if* S has a second-grade sensation, then S cannot refrain from issuing a *c*-judgment. And as before, the reason has to do with optimal design: God deemed this set-up best for the preservation of our health.

²²If *c*-judgments are beliefs, however, then the result is reminiscent of a Reidian dual-component theory of perceptual experience, in which perceptual experience contains as a constituent doxastic states.

D4 God brought about E for the purpose of making S falsely c-judge that bodies are p-colored.

D4 corresponds to clause (iv) of DECEPTION. By ‘brought about’ we mean that God is broadly causally responsible for the makeup of our sensory system, which includes the way things look. We see three ways in which D4 might be false. Each objection targets clause (iv) by denying that God is responsible for our deception, but they do so differently. The first objection trades on our finitude; God designed us and our sensory systems, but without any intent *per se* to deceive us. Bodies look p-colored because of a design glitch. On the second objection, *we ourselves* generate the misleading evidence via our habitual violations of epistemic norms. On the third objection, given our violation of epistemic norms, while S does not *generate* the evidence, *S misunderstands* what the second-grade is telling her. And it’s only because of this misunderstanding that S is deceived; once the norm is followed, S realizes that the second-grade isn’t misleading after all. The common thread of these objections is that although S can be deceived, it’s not due to God. Responding to these objections substantiates SENSORY DECEPTION.

Objection One. The first objection to D4 leverages the Sixth Meditation response to the problem of dropsy (AT VII: 87-9/CSM II: 60-1). The thirst sensations S has when suffering from dropsy are both deceptive and harmful. Descartes responds by noting that God’s intention in designing our system was to give us the best of all possible sensations (Sixth Meditation, AT VII: 87-8/CSM II: 60). However, this optimal design is implemented in finite mind-body composites. And finitude allows for glitches in our sensory system. An aberrant motion in the body can cause a sensation that is harmful to us. Such glitches are due to our finitude, not God’s design of our sensory system (Sixth Meditation AT VII: 88-9/CSM II: 61). By extension, the thought then is that p-color sensations misrepresent bodies as p-colored due to our finitude.

This objection fails. Unlike dropsy, p-color sensations are *systematically* misleading. Dropsy is a rare or one-off affair, explainable by our finitude and the failure of proper correspondence between physiological stimulation, brain impression, and sensation. By contrast, the error inherent to p-color sensations leaves every impression of design; bananas standardly look p-yellow (but are not), leaves standardly look p-green (but are not), and they all do so in predictable circumstances.

If I want to know what an agent’s intentions were in creating some system, then I should examine how it behaves under optimal conditions, because it’s those conditions that reveal what the agent wanted. The fact that dropsy rarely occurs, and only occurs when the body isn’t functioning properly, tells against God’s bringing about dropsy so that S would drink. We’ll return to this point later (Section 3.2), but the point for now is that we have precisely

the opposite situation in the case at hand: the systematic illusoriness of color sensations is not a glitch. Things are working how they are supposed to be working. Our visual system is *supposed* to falsely attribute p-colors to bodies. It's reasonable to infer, then, that it was a part of God's design plan.

Objection Two. The second objection leverages the Fourth Meditation account for how to avoid intellectual error: *the truth rule*. This rule says that a subject S, should not judge that *p* unless S is occurrently clearly and distinctly perceiving that *p*. (Third Meditation, AT VII: 35/CSM II: 24; Fourth Meditation, AT VII: 60-1/CSM II: 41) Since p-color sensations are not clear and distinct, when S affirms them she violates the truth rule. I improperly use my will by extending it to that which is obscure and confused. Therefore, the source of the error lies in us, not God. So, God is not a deceiver.

The truth rule is restricted to r-judgments. For it to have any teeth, it must apply to judgments that we can revise or cease to have. So it cannot apply to c-judgments. But it can still be apt, and relevant to D4, if appropriately supplemented. The supplementation draws a connection between judgments we can revise—i.e. r-judgments—and how bodies look at the second-grade. This is critical, because even if we do violate the truth rule, we violate it because of the misleading phenomenological evidence at the second-grade—evidence, that has, on the face of it, nothing to do with us. So what we really need is some combined principle, to the effect that bodies will look p-colored only when S fails to exercise her will properly—that is, only when S r-judges in a way that violates the truth rule.

Such a combined principle says that when S judges *in a way where she could have judged otherwise*, she creates the misleading appearance of bodies looking p-colored. The misleading phenomenological evidence at the second-grade is due to S in the first place.

Nolan (2011: 98) endorses something like this move in his attempt to acquit God of deception. Recall that Nolan reads Descartes as a color nominalist, specifically a form of strong eliminativism where color appearances are not sensory or perceptual, but wholly due to habitual (false) judgments. Nolan's motivations here are multi-faceted, but the key point for present purposes is that if misleading p-color appearances are due to false judgments—specifically, false habitual, and revisable judgments—then God, it would seem, is off the hook. Nolan has us ask why our senses *seem* to teach us that p-tastes and p-colors are in bodies—why does this seem to be a “teaching of nature”?—when in fact they aren't, and when in fact our senses don't even *purport* to teach such a thing (Ibid., 97). In speaking of how things “seem”, Nolan is asking, of Descartes, why the phenomenological evidence is how it is. And Descartes' answer, according to Nolan, is that we *generate* these misleading seemings (i.e. misleading phenomenological evidence E) via engrained-since-childhood, habitual judgments that violate epistemic norms. In other words: we mislead ourselves. Indeed, this self-deception

runs quite deep, for strictly speaking, we do not even have (second-grade) sensations *of* p-color, as such (Ibid., 100). We merely think we do.

Of course, we concur that judgments play a role in fixing sensory phenomenology. But such judgments cannot be of the revisable, governed-by-epistemic-norms sort that Nolan envisions. And as Nolan himself notes, it's only those judgments that acquit God of deception.

Phenomenology-generating judgments cannot be of the revisable, governed-by-epistemic-norms sort for this reason: bodies *still* look p-colored when upon believing eliminativism and applying the truth rule, I cease judging that bodies are p-colored. To see this, compare:

(6) The book looks (seems) p-red to S.

(7) It looks (seems) to S like the kitchen was ransacked.

'Looks'-statements (or 'seems'-statements) are utterances of sentences like (6) and (7). But the sense of looks in (6) is quite different from that in (7). Sentence (6) employs a *non-comparative phenomenal sense* of 'looks' as we saw in (4) and (5) above, whereas (7) employs an *epistemic sense* of 'looks' (e.g., Chisholm 1957). When 'look' is used epistemically, the sentence conveys what is subjectively probable, conditional on the evidence available. A mark of epistemic looks is that they go away in the presence of a defeater if the agent is rational. So, if you are rational, it will stop looking as if a hungry burglar ransacked your kitchen once you see the paw prints and deduce that it was a raccoon.

That's not so with when 'look' is used phenomenally. Even when S learns that she is suffering an illusion, and that the book is not p-red, she will still sincerely utter (6). This point is not limited to p-color:

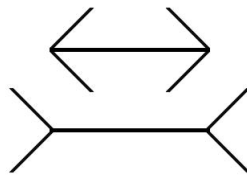


Fig. 1: The Müller-Lyer Illusion

The second line does not stop looking longer than the first even when I know that they are equal in length. But now notice: relevant for this new principle is the phenomenal sense of 'looks'. Just like the lines don't cease (phenomenally) looking to be different in length, bodies will not cease looking p-colored when the truth rule is applied. Surely, Descartes knew this.²³

²³Müller-Lyer type cases—or *known-illusions*—are compatible with the present story about c-judgments. When S denies that the book is p-red post-*Meditations*, she is expressing her *lack of endorsement* for her

Objection Three. Instead of claiming that S generates E, perhaps the worry is that, given S's violation of another epistemic norm, S *misunderstands* E. Insofar as S is deceived by sensations, the mistake is on S, not God.

Simmons (1999) advocates this sort of move. Whether a sensation will give rise to a true or a false r-judgment depends on what S believes about the *function* of the sensation. Given their biological function, the idea is that sensations are not designed to guide us in metaphysical inquiry. As Descartes says, "I misuse them [sensations] by treating them as reliable touchstones [*regulis certis*] for immediate judgments [*immediate dignoscendum*] about the essential nature of the bodies located outside us" (Sixth Meditation, AT VII: 83/CSM II: 57-8). Or as Simmons puts it: "[w]hat is important here is that the level at which sensations get their representational hook on the world is the level of ecology not physics" (1999: 356). If S understands the proper function of sensations, then sensations will guide S to making true judgments about bodies in her environment. This is the sense in which sensations are *materially true* for Simmons (1999: 363). However, if we misunderstand the function of sensations—believing that they are accurate guides to the metaphysics of *res extensa*—then we will make false r-judgments about bodies. This is the sense in which sensations are *materially false* for Simmons (1999: 363; cf. Nelson 1997). In short, Simmons is providing another epistemic norm that is supposed to guide our r-judgments. Call this *the function rule*: S ought to clearly and distinctly perceive the essential biological function of a given sensation before making any r-judgments based on that sensation.

The function rule, as a species of the truth rule, is restricted to r-judgments. So at least one portion of this objection trades on the assumption that it is r-judgments alone that matter for being deceived, not c-judgments; for all the objection says, our c-judgments may still be false. This is an important issue, which we address next in discussing D6.²⁴ The more immediate issue at stake, however, is this: does our failing to follow the function rule not just imply that we misunderstand E, but that we misunderstand E in a way that shows that E is not really deceptive?

It's unlikely. For a sensation to represent, it must make a 'claim' about the world. For Simmons, these claims are not metaphysical but "ecological". Let's grant that. Here's an ecological content: <bananas are edible>. It is ecological in Simmons' sense since it is the sort of claim which, when taken as the content of a sensation, promotes health and fitness. It's perhaps not a metaphysical claim, insofar as it doesn't tell us about the *deep* or *essential*

c-judgment that says otherwise. This is not surprising given the non-inferential nature of c-judgments, and our more general recognition of having cognitive states that we fail to endorse (Quilty-Dunn 2015). See also Section 2.4.

²⁴Simmons (1999) employed this move before her introduction of c-judgments (2003). Our guess is that she would now reject D6.

nature of bananas, relative to the proper modes of *res extensa*, and independent of us as agents in the corporeal world. Indeed, the sensory representation of such metaphysical claims may (as Simmons notes) even be a hindrance to our survival. But the ecological claim is (as it happens) a true claim; it states a fact about bananas. If we followed the function rule, then, we would realize that our sensations make claims of *this* sort, instead of claims about what bodies are like as conceived by the Cartesian physicist (Simmons 1999: 353, 355).

The problem, though, is that whatever exactly ecological contents look like, they will include p-colors. This is for the simple reason that p-colors are highly ecologically-relevant, even if we do not want to say that p-colors are ecological properties strictly speaking (like *edibility*). To know which bananas are edible, or at least which are most nutritious, I need to know which are ripe, which means I need to know which look p-yellow. As such, it stands to reason that amongst the ecologically-oriented contents will be things like <bananas are p-yellow>, since the banana's looking p-yellow just consists in the representation of this content. And while unlike <bananas are edible>, <bananas are p-yellow> is false, our sensations tell us otherwise by virtue of their attributive structure, and so are deceptive on that count. Schematically, they say something like <x is F>, and so say *of* bananas ('x') that they are p-yellow ('F'). Indeed, this attributive structure is critical to sensations' ecologically-oriented role. Claims do not have to assert metaphysical, essential, or fundamental truths to be misleading.

So the distinction between metaphysical and ecological claims doesn't move the needle. When we heed the function rule, and learn that our sensations 'only' make ecological claims, we are not thereby coming to realize that our sensations reserve themselves to *true* ecological claims about bodies (as opposed to false metaphysical claims), and thus are not misleading. Simmons might insist that <bananas are p-yellow> is true. However, not only would the weight of the evidence be against her from Descartes' remarks on p-colors alone, but the more basic point is that ecological contents don't have to be true to serve their survival-promoting role, and it's question-begging to assume others. We are not any worse off, *qua* survival, for bodies not being p-colored. What matters to finding ripe bananas is that they *look* p-yellow. And to look p-yellow, a banana doesn't have to *be* p-yellow. It only need be represented as p-yellow.²⁵

²⁵When saying that our sensations are materially true, Descartes actually only tells us that his "senses report the truth much more frequently than not" (Sixth Meditation, AT VII: 89/CSM II: 61)". The "much more frequently" line is consistent with saying that our senses misrepresent some ecologically-relevant properties (like p-colors) as properties of bodies, even systematically so.

2.4 Regarding D6

Recall:

- D6 If (i) S judges falsely that p-bodies are p-colored in virtue of E [D3], (ii) God brought about E for the purpose of making S falsely judge that bodies are p-colored [D4], and (iii) God knows that bodies are not p-colored [D5], then God is a deceiver.

It might seem like D6 isn't in need of any defense, as the consequent seems to follow given DECEPTION. This is the line Patterson (2016: 98) takes, emphasizing as she does that if God gave us a natural propensity to make false c-judgments, then God's being a deceiver is a genuine problem. Yet this inference can be blocked by arguing that DECEPTION is only a sufficient condition for deception if its use of 'judgment' concerns *r-judgments*. And since r-judgments *are* revisable and governed by epistemic norms like the truth and function rules—rules S flouts—God is off the hook. If all goes well post-*Meditations*, then S no longer *r-judges* that bodies are p-colored.

The objection can be fleshed out by considering again how the truth and function rules were employed in the discussion of D4. There, the former was used—supplemented by the idea that bodies look p-colored due to a failure to properly exercise the will in r-judgements—to examine how E might be self-manufactured, while the latter was used to show how we might somehow be mistaken about E. But suppose instead we grant that God *did* bring about E and that E is deceptive—i.e. suppose that bodies look p-colored despite not being so, and not because of us. One could grant this while denying that God brought about E for the purpose of us making *false* judgments. This would not work if we meant c-judgments, since we will always make c-judgments come what may, and given eliminativism, they will always be false. So this cannot be an objection to D4 *per se*. In this way, on the intended reading, the objector can grant that D4 is true. However, if she insists that it's only r-judgments that matter for deception, then the truth and function rules can be re-envisioned simply as guides for what r-judgments God *did* intend for S to form, irrespective of what S c-judges.²⁶ So while D4 is true for c-judgments, it does not follow that God is a deceiver. Thus D6 is false.

²⁶Pre-*Meditations*, our r-judgments about p-color are false. Strictly speaking, more than the truth and function rules are needed to show that God is not a deceiver pre-*Meditations*, however. We also need to assume that deception must be intentional. For if there could be deception without an intent to deceive (recall fn. 5), one could simply say that we are deceived by God with respect to our r-judgments because we falsely r-judge due to E, even though God only intended for us to form *true* r-judgments, something we would have known once we read the *Meditations*.

Yet it's doubtful that letting 'judgment' refer to c-judgments renders DECEPTION false. If the idea is that c-judgments, at best, show that the sensory system is deceived and not S, then this worry is blocked by the fact that c-judgments are issued by the will—an active, not passive part of the mind-body composite. And while we have been neutral on whether c-judgments are beliefs, they certainly seem belief-like enough for deception. They are semantically evaluable as true and false; indeed, as was noted, a judgment that p —whether a r-judgment or c-judgment—will always involve putting p before one's mind *as true*. We might imagine a scenario in which, upon seeming to see a barn in the distance, I momentarily c-judge 'There's a barn'. Perhaps this judgment doesn't get as much cognitive traction as a belief (assuming the c-judgment is not a belief), but it's not cognitively idle, since it can influence my behavior. (I'll be more apt to approach the barn c-judging that it's before me.) Yet if I learn that a trickster just put up a barn facade, then I was deceived, despite only (falsely) c-judging. So c-judgments are perfectly apt to figure into an account of deception.

However, there is certainly *something* curious happening post-*Meditations*, and perhaps this is what gives rise to suspicions about D6. Suppose S has yet to read the *Meditations*. S has all sorts of false r-judgments: these are the r-judgments that are involved in central planning and conscious inference. Amongst these false r-judgments will be *that bodies are p-colored*. Like all r-judgments, this is formed from S's epistemic habits, and God cannot be impugned by them. But then S reads the *Meditations*. S will no longer r-judge that bodies are p-colored, assuming all goes well. And this will impact S's behavior. If you ask S whether bananas are p-yellow, or whether S would bet a large sum of money that the p-color of healthy grass is p-green, S would say 'no'. So the worry is that, given S's new cognitive and behavioral profile, S is now not deceived about bodies being p-colored, false c-judgments notwithstanding. R-judgments take priority when assessing the epistemic status of an agent. And you simply cannot be deceived while believing (or r-judging) the truth.

But you can. Take implicit bias. Subjects profess to not have a corresponding *explicit* belief that p , despite evidence that they have the relevant implicit attitude. Likewise, post-*Meditations*, S will explicitly deny that bodies are p-colored, even though her will issues a claim (a c-judgment) to the effect that they are, and even though this issuance is irrevocable.

Now it's often thought that implicit bias is caused by some associative process/associative structure. This is usually glossed as some type of *evaluative* association, such as an association between a valence (e.g., negative affect) and a concept (e.g. BLACK MALE). Eric Mandelbaum (2016) calls this thesis *associative implicit bias* ('AIB'), and convincingly argues that AIB is false. In its place, he proposes *The Structured Belief* hypothesis. Here, implicit bias is underwritten by unconscious *beliefs*. These beliefs are not mere associations. They are propositionally structured mental representations to which S bears the belief rela-

tion. So instead of maintaining, to use Mandelbaum’s example, that implicit racists merely associate (say) BLACK MALE and DANGEROUS, the hypothesis is that implicit racists have a belief with the structure BLACK MALES ARE DANGEROUS.

Suppose then that one forms the unconscious belief *that black males are dangerous* via deceptive means. When asked, you don’t endorse this belief. Yet you act in ways that suggest you believe this. For example, when on a subway and a black male sits next to you, you move away. When you walk to your car at night, and you see a black male behind you, you quicken your pace. Such examples are emblematic of the behavioral patterns of those who exhibit implicit racial biases. The upshot is that one’s behavior is not merely a function of what one expressly avows, or what one consciously believes.

Whatever c-judgments are—either full-blown beliefs with muted dispositions or belief-like judgments—they are not mere associations. We can thus substitute Mandelbaum’s implicit beliefs with Descartes’ will-issued c-judgments, and consider our subject once again. S’s false c-judgment is made automatically and is at least confusedly presented to consciousness (Simmons 2003: 554 fn. 13)—much like the implicit racist who unconsciously, and falsely, believes that black males are dangerous. Upon reading the *Meditations*, S does not simply no longer endorse these judgments. S endorses eliminativism and so believes that bodies are not p-colored. And this belief is true. Hence S at once c-judges (without endorsing) that p , while r-judging that $\neg p$, much like classic cases of implicit bias.²⁷ The implicit racist’s unconscious belief need not be formed via deceptive means, but they can be, and when they are, we have no trouble saying the implicit racist was deceived irrespective of what she expressly avows. Her behavior—moving seats on the subway—is then in turn also the by-product of deceptive information. Similarly, S’s c-judgment is occasioned by misrepresentation at the second-grade, and will show up in her behavior, e.g. in how she interacts with bodies. And this all holds irrespective of whether S’s r-judgments are true. The requirement that DECEPTION can only be framed in terms of r-judgments is thus too strong.

3 Don’t Panic

It wouldn’t be ideal if granting the foregoing argument from SENSORY DECEPTION simply revealed that Descartes is inconsistent and perhaps confused. So the question is whether there

²⁷Implicit bias need not be a case of having an unconscious belief that p while refraining from consciously endorsing p ; it can also be a case of unconsciously believing that p while consciously believing that $\neg p$. Bendaña and Mandelbaum (2021) call the latter *belief fragmentation*. This allows for inconsistent information to be selectively accessed such that there is no “interfragment consistency”—that is, no fragment (roughly, a data structure) stores both p and $\neg p$ at once (Ibid.). We contend that post-*Meditations* the subject is suffering from something like belief or belief-judgment fragmentation.

is a way for Descartes to accept the argument on his own terms. Patterson (2016: 100-102) says there isn't. We disagree.²⁸ Although there are no texts where Descartes unambiguously addresses this issue, Descartes does offer the *resources* for a solution where we can understand in what sense God can be a deceiver without upending his philosophical system. The solution isn't a sensory theodicy by which one tries to explain away the deception. Rather, Descartes has within reach a theodicy in which he can bite the bullet: sensations are deceptive and God is a deceiver on their account, albeit not problematically so.

Setting out such a solution may seem impossible, for there are too many passages where Descartes claims that God cannot be a deceiver. We cannot examine all of these passages here. But consider this paradigmatic one:

It is impossible [*non posse*] that God should ever deceive me [*me unquam fallat*]. For in every case of trickery or deception [*fallacia vel deception*] some imperfection is to be found; and although the ability to deceive appears to be an indication of cleverness or power, *the will to deceive is undoubtedly evidence of malice or weakness* [*malitiam vel imbecillitatem*], and so cannot apply to God. (Fourth Meditation, AT VII: 53/CSM II: 37; cf. *Principles* I.29, AT VIII:16/CSM I: 203; emphasis added).

Here, Descartes claims that every form of deception is incompatible with God's perfect nature. Why is that the case? The reasoning in this passage (and similar passages) runs as follows. Call this the *incompatibility argument*:

- I1 Every form of deception involves malice (one kind of imperfection) or weakness (another kind of imperfection).
 - I2 Malice is incompatible with God's benevolence (a divine perfection).
 - I3 Weakness is incompatible with God's omnipotence (a divine perfection).
- ∴ God cannot be a deceiver.

While (I2) and (I3) are implicit in the above passage, the Incompatibility Argument is the theoretical grounds for denying that God is a deceiver.

However, we *do* argue that the Incompatibility Argument, and the passages that bear it out, do not tell the entire story about Descartes' views on divine deception. Take (I1). To deny (I1), Descartes would have to allow for:

²⁸Patterson denies the viability of any Malebranchian solution that says "it is fine for the senses to dispose us to false judgements... since the senses aim at survival rather than truth" (2016: 101). Patterson contends that Descartes can't go in for this line because he nowhere claims that false c-judgments are designed to preserve our health (Ibid. 105). We disagree: the point isn't that *being false per se* is relevant for health-preservation, but that it's *no bar* for health-preservation.

OMNI-COMPATIBLE DECEPTION: God’s ϕ ’ing is an act of omni-compatible deception if God’s ϕ ’ing is an act of deception that (i) does not involve malice (incompatible with omnibenevolence) and (ii) does not involve weakness (incompatible with omnipotence).²⁹

We think Descartes can allow for omni-compatible deception, and indeed there are passages, if only *glimmers*, that point precisely this way. We’ll focus mostly on clause (i), given that these passages only explicitly reference malice, and because it’s not obvious what Descartes means by ‘weakness’ here anyway. On at least one immediately plausible reading, “malice or weakness” expresses not a true disjunction, but close synonyms. This would make weakness something like *moral weakness*, which is functionally equivalent to malice, thus rendering (I3) and clause (ii) in OMNI-COMPATIBLE DECEPTION superfluous. Alternatively, weakness could be tied to omnipotence, with the idea being that surely God could have achieved his ends (whatever they are) in some *other way* that did not involve deception. This is an interesting issue, although one that is not taken up here for sake of space.³⁰

Are these passages anomalous? In a way they are, but in another way they are not. Given the preceding arguments, there is now a meta-interpretive pressure to look at them anew. At a first-order interpretive level—looking at these passages independently from our argument from SENSORY DECEPTION—such a dismissal is more warranted, especially in light of Descartes’ denials of God’s deceptive nature elsewhere. However, when they are read against the background of our argument, we now have new reasons (e.g. those driving from consistency) to take otherwise obscure passages more seriously.

We’ll begin by examining passages where Descartes discusses the possibility of omni-compatible deception. We’ll then turn to maliciousness, and consider two questions: (1) is God malicious in giving S deceptive second-grade sensations? And (2): Is God malicious in making it such that S cannot help but issue false c-judgments in response to the second-grade? The answer to both is *no*. Finally, we argue that omni-compatible deception does not upend other aspects of Descartes’ system.

3.1 Descartes on Omni-Compatible Deception

There are two main texts where Descartes acknowledges the possibility of divine deception. First, in *Objections and Replies* II Mersenne writes:

You say that God cannot lie or deceive. Yet there are some schoolmen who say

²⁹We set aside features that are incompatible with omniscience, as Descartes does not mention anything related to omniscience in the Fourth Meditation passage above.

³⁰Though see fn. 33.

he can. Gabriel...and Ariminensis among others, think that in the absolute sense God does lie [*mentiri*], that is, communicate to men things which are opposed to his intentions and decrees. Thus he unconditionally said to the people of Nineveh, through the prophet, ‘Yet Forty days and Nineveh shall be destroyed.’ And he said many other things which certainly did not occur, because he did not mean his words to correspond to his intentions or decrees...Cannot God treat men as a doctor treats the sick, or a father his children? In both cases there is frequent deception [*saepe decipiunt*] though it is always employed beneficially and with wisdom [*sapientur & cum utilitate*]. For if God were to show us the pure truth, what eye, what mental vision, could endure it? (AT VII: 125-6/CSM II: 90).³¹

In pointing out false texts in scripture that may lead to false judgments, Mersenne presses on Descartes’ seemingly unrestricted contention that God cannot be a deceiver. Mersenne offers the doctor-patient example to explain how deception could be permissible for God. A doctor might say something false to a patient about their sickness with the intent of helping them. This isn’t obviously problematic. Similarly, there might be falsities in scripture that are beneficial for believers. Perhaps then such cases of deception would be permissible for God because it doesn’t conflict with his benevolence. Now notice how Descartes responds:

I would not want to criticize [*nolim tamen reprehendere*] those who allow that through the mouths of the prophets God can produce verbal untruths which, like the lies of doctors who deceive their patients in order to cure them, are free of any malicious intent to deceive [*desit omnis malitia deceptionis*]. (AT VII: 143/CSM II: 102).

While Descartes does not explicitly endorse non-malicious deception, that he claims that he will not criticize the possibility of the doctor-patient case suggests an implicit endorsement. Throughout the *Objections and Replies*, Descartes is quite happy to point out when his interlocutors fall into contradictory positions. If it were in fact contradictory for God to deceive full stop, then Descartes would presumably point that out. One part of his response is particularly telling: he claims that doctors can deceive without having “any malicious intent to deceive.” This stands as a qualification to (I1).

Second, Hobbes raises the same query:

The standard view is that doctors are not at fault if they deceive their patients for their health’s sake, and that fathers are not at fault if they deceive their

³¹Note that *mentiri* is translated by Cottingham et. al. as “lie.” But a more suitable translation is “deceive” given the broader context of this passage. For a discussion of the medieval accounts of deception referred to here, see Perler (2010).

children for their own good. For the crime of deception consists not in the falsity of what is said but in the harm [*injuria*] done by the deceiver. (*Objections III*, AT VII: 195/CSM II: 136)

Descartes responds like this:

[I don't] require that we can in no case be deceived (indeed, I have readily admitted that we are often deceived). All that I require is that we are not deceived in cases where our going wrong [*noster error*] would suggest an intention to deceive on the part of God [*voluntatem in Deo testaretur*]; for it is self-contradictory that God should have such an intention. (*Third Replies*, AT VII: 195/CSM II: 136-7)

This response is interesting, if only because Descartes doesn't answer Hobbes' question. The doctor's intention is to aid the patient; the deception is a means to that end. So, again, we ask: if Descartes wanted to definitively rule something out, why did he not do so?³²

Descartes is being cagey. He is well-aware of the places where he claims that God cannot be a deceiver, and the crucial role that this claim plays in his metaphysics and epistemology. Descartes might have feared, then, that he would upend his entire system, and thus did not want to make a full concession to Mersenne and Hobbes. Yet our contention is that Descartes failed to see two key things: first, that his prior commitments concerning the structure of our sensory systems commits him to God being a deceiver in some sense, and second, that in fact he can admit this without wreaking havoc on his philosophical system.

Suspending judgment regarding whether this is what Descartes actually believed or meant, here is our suggestion for what Descartes *could*—and indeed in light of our preceding argument from SENSORY DECEPTION—*should* have meant: some acts of deception are not malicious. That is, they are not meant to, nor cause, harm. And thus such acts of deception are omni-compatible insofar as they are compatible with God's benevolence.

Equating non-maliciousness with non-harmfulness is sensible given the standard meaning of 'malice' but also Hobbes' point that the "crime of deception consists...in [its] harm." Harm [*nociturus*] is often bodily, as in Descartes' Sixth Meditation discussion of dropsy (AT VII: 85/CSM II: 59). But other harms are non-bodily. In the case of the evil deceiver of the First Meditation, the harm is plausibly *psychological* (AT VII: 23/CSM I: 15). This is evidenced by the kinds of passions—such as dread and fear—expressed by the meditator when she considers the possibility that she could be deceived about all matters (see Section 3.3).

Now there are two components of deception that are relevant for present concerns. First, there are the *deceptive* appearances at the second-grade, i.e. sensory evidence E. Second, there are the false c-judgments that we issue in response to E. Let God be a malicious

³²Robinson (2013: 522) also reads Descartes as granting omni-compatible deception.

deceiver if and only if either of these components harms the subject who is deceived. So, the question we must now answer is this: is the deception inherent in our sensory systems malicious?

3.2 Benefit and Harm in Sensory Systems

Ascertaining whether God’s designing of E makes him malicious can be clarified by considering two claims. First: the biological function of sensations is to preserve the health of the mind-body composite (Sixth Meditation, AT VII: 83/CSM II: 57). Second: God gave us the best of all possible sensations (Sixth Meditation, AT VII: 87-8/CSM II: 60).

The conjunction of these claims implies that God gave us the best sensations *qua* their functional role—“the one sensation which, of all possible sensations [*ex omnibus quos inferre potest*], is most especially and most frequently conducive to the preservation of the healthy man” (Sixth Meditation, AT VII: 87-8/CSM II: 60). So the tighter gloss on the current question is this: assuming God would be malicious if he gave us anything less than the best of all possible sensations, is the biological function of p-color sensations compatible with God making them systematically misleading?

We won’t defend the claim that the misleading nature of p-color sensations somehow makes them ‘the best’ *qua* their aiding us in, e.g. finding food. All that need be the case is that p-color sensations being misleading is *no bar* to their optimally playing this sort of role.

When asked how our various sensations could play their part in the preservation of the mind-body composite, whether they misrepresent is irrelevant. What matters is *reliability*. Mendelovici (2013) points out that a test can fail to be valid while still being reliable. For example, the SAT may fail to be valid insofar as it fails to detect what it’s intended to detect—e.g. predicting success in college (Ibid.: 422). Still, it could be reliable insofar it tends to yield the same results in different contexts. Similarly, while some representations might be veridical and reliable, others may not. Our interest is in those that *reliably misrepresent*: reliable, in that they respond similarly in similar circumstances; and non-veridical, since the world does not have the properties they represent it to have. We want, as Mendelovici puts it, a representation that “[gets] things wrong in the same way all the time” (Ibid.: 423).

P-color sensations are reliable misrepresentations. They are illusory, but in a systematic way. Paraphrasing somewhat, Mendelovici (Ibid.: 423) tells us that, for any (indicative) mental representation R and any property F, R reliably misrepresents F if and only if: (i) R represents objects as having F; (ii) most or all of the relevant objects don’t have F; and (iii) tokens of R do or would non-veridically represent objects as having F in the same types of circumstances on separate occasions. These conditions are sufficient to reconcile God’s

designing us to have systematically misleading p-color sensations despite having a health-preserving functional role and their being optimal in fulfilling that role. Here’s why. Suppose S only wants to identify ripe bananas. Ripe bananas look p-yellow. By eliminativism, they are not p-yellow. So, when S sees bananas, she represents them as having a property they don’t have. But that’s not a problem insofar as finding food goes. What *would* be a problem is if S’s sensation represented the banana as p-yellow in context C, p-green in context C*, p-magenta in C**, and so on, where C, C*, and C** are all tokens of the same context-type. In all contexts, S’s sensation would misrepresent. But it wouldn’t be doing so in a reliable manner. And that would prevent S from finding food in an optimal manner, since it would sever the connection between *being ripe* and *looking p-yellow*.

Considerations that hold at the second-grade can be extended to the third-grade. C-judgments are part of our sensory system and function to preserve health. Just as the non-veridicality of p-color sensations is irrelevant on this score, the falsity of c-judgments is too. What matters is the properties they represent bodies as having, and how they do this.

Notice too that deception *per se* need not be beneficial. In the patient-doctor example, what is benefiting the patient is the doctor’s *speech act*—her utterance of ‘you are going to get better.’ That claim might be false, and the doctor might know it. Since the doctor wants her patient to believe that she will get better, the doctor deceives her patient. Still, the speech act isn’t beneficial in virtue of the statement it expresses being false. After all, the speech act would *still* be beneficial even if it were true. When a doctor says to a patient ‘you are going to get better,’ the reason this helps (or so we are told) is that it aids the patient psychologically. These effects would still result even if it were true that the patient was going to get better. (Maybe the doctor was right to say this even though she didn’t know it.) Similarly, p-color sensations are not beneficial in virtue of being deceptive, but in virtue of the way they attribute p-colors to bodies. The deception is simply *non-malicious*.³³

3.3 Systematic Concerns

The veridicality of clear and distinct perceptions, the resolution of the dream doubt, and the proof of the external world rely on *some* anti-deception premise. Yet for each all that’s required is an anti-*malicious* deception premise.

A core claim of the *Meditations* is that clear and distinct perceptions are veridical because

³³Question: *why* does God deceives us, even if only non-maliciously? If deception *per se* isn’t beneficial, it’s natural to wonder why God didn’t simply make bodies p-colored, and let us reap the benefit anyway. That might be a sign of weakness, which if seen as incompatible with God’s omnipotence, constitutes a distinct worry. The cogency of this question assumes that God could have made p-color a property of *res extensa* without making p-color sensations any less beneficial. We doubt there is any relevant sense in which God ‘could’ have done such a thing (Nelson & Cuning 1999), but this is not a point discussed here.

God is not a deceiver. Yet if God were to deceive S about her clear and distinct perceptions, this would be malicious deception, which is exactly *why* Descartes says that it's inconsistent with God's nature for him to deceive us about our clear and distinct perceptions. One reason intellectual deception would be malicious is that it would prevent S from attaining virtue and ultimately happiness. And for God to prevent S from attaining happiness would be to harm her. For Descartes, genuine happiness—the supreme satisfaction or contentment of mind—results from being virtuous (see, e.g., Letter to Queen Christina 20 Nov. 1647, AT V: 83/CSMK: 325). Cartesian virtue consists in the firm and constant resolution to use the will well. Here, using the will well consists in practically judging and acting according to knowledge of the truth (Letter to Princess Elizabeth 15 Sept. 1645, AT IV: 291/CSMK: 266). Knowledge of the truth is a necessary (but not sufficient) condition for virtue, and broadly construed consists of four truths: the existence of God, the distinction between mind and body, the immensity of the universe, and the interconnectedness of the universe (Letter to Princess Elizabeth 15 September 1645, AT IV: 291-3/CSMK: 265-6). Focus on the former two. According to Descartes, S needs to know that God exists and that everything in existence is a manifestation of God's will, and S must also know that the mind is distinct and more noble than the body in order to regulate our passions and exercise virtue (Ibid.). These two metaphysical truths are grounded in clear and distinct perceptions (Parvizian 2021).

If S did not know these truths, Descartes would say it's impossible for S to exercise virtue, and thus she would be incapable of acquiring virtue and happiness, at least in their ideal forms (Dedicatory Letter to Elizabeth, AT VIIIA: 3/CSM I: 191). This implies that if God were to deceive S about, say, the real distinction between mind and body, then he would be preventing S from acquiring ideal virtue and happiness—a clear harm. As such, intellectual deception is inconsistent with God's benevolent nature.

What of the dream doubt, and its extensions regarding our apparent lack of epistemic certainty in the existence of external bodies? Here again, that God is a non-malicious deceiver causes no problem. For if God were to deceive S about the *existence* of bodies, He *would* be harming S, for she would be out of touch with reality, which would cause psychological harm.³⁴ The psychological harm induced by not being in touch with reality is brought out at the end of the First Meditation, where the meditator expresses “dread” and “fear” in the face of the “inextricable darkness of the problems I have now raised” (First Meditation, AT VII: 23/CSM II: 15). If Descartes could not resolve the dream doubt, much of these passions and

³⁴Notice that if the dream doubt cannot be resolved, then we risk not being in touch with reality *at all*. By contrast, p-color sensations are illusions, and it's a conceptual truth that illusions (unlike hallucinations) make us aware of the objects they represent.

thoughts would persist, yielding a psychologically and practically harmful state of affairs.

Finally, Descartes' proof of the external world is handled similarly. The proof is a disjunctive syllogism (Stuart 1986: 22). Either my ideas of material objects are caused by (i) a unknown faculty of the mind, (ii) God, or (iii) by actual material objects. Yet my ideas of material objects are neither caused by (i) or (ii). Hence, they must be caused by (iii) actual material objects. Now the sub-argument against (ii) appeals to God's non-deceptiveness (Sixth Meditation, AT VII: 79-80/CSM II: 55), but all that's relevant here when it comes to the argument's validity is God's non-deceptiveness *with respect to the existence of external bodies*. As we have seen, deceiving us in that respect is malicious.

4 A Concluding Irony

Descartes is committed to SENSORY DECEPTION, and so, given facts about our sensory system, Descartes' God is a deceiver. While Descartes' broader system remains intact, we should rethink how to approach that system.

The *Meditations* contains two theodicies: the epistemic theodicy of the Fourth Meditation and the sensory theodicy of the Sixth Meditation. In the Fourth Meditation, Descartes needs to resolve the epistemic worry that God might deceive us in our clear and distinct perceptions, which would preclude us from acquiring *scientia*. Having completed the Fourth Meditation, it's tempting to attribute to Descartes the view that God can never be responsible for our forming *any* kind of false judgments. The problem is that this makes a consistent sensory theodicy in the Sixth Meditation impossible. For in the *Sixth Replies*, Descartes clarifies his account of our sensory systems so as to make clear that they include c-judgments: will-issued, irrevocable, false sensory judgments. However, if we see Descartes as being primarily concerned with *benevolence*, the tension dissipates. For now the Fourth Meditation in fact teaches us that God cannot be responsible for us forming any *harmful* judgments. That's bedrock. And since *false* r-judgments about clear and distinct perceptions *would* be harmful, they cannot be due to God. However, the standards for harm in the sensory theodicy of the Sixth Meditation are different. C-judgments are not harmful on account of their being false, but they are beneficial for our health in virtue of the way they represent bodies. So God, in causing us to have c-judgments, is both a deceiver *and* benevolent. What's ironic is that it takes realizing that God is a deceiver to properly see the theoretical centrality afforded to benevolence within Descartes' system.³⁵

³⁵For helpful comments and discussion, we thank an audience at the Chicago Early Modern Roundtable, Margaret Atherton, Devin Sanchez Curry, David Hilbert, Kristen Irwin, Alan Nelson, Larry Nolan, John Whipple, and the anonymous referees at *Synthese*.

References

- [1] Atherton, M., (2004). Green is like bread. In Ralph Schumacher (ed.), *Perception and Reality: From Descartes to the Present*. Mentis.
- [2] Averill, E.W. (2005). Toward a Projectivist Account of Color. *Journal of Philosophy* 102 (5):217-234.
- [3] Averill, E.W. & Hazlett, A. (2011). Color objectivism and color projectivism. *Philosophical Psychology* 24 (6):751 - 765.
- [4] Bendaña, J. & Mandelbaum, E. The Fragmentation of Belief. (2021). In *The Fragmentation of Mind* (eds. Dirk Kindermann, Cristina Borgoni, & Andrea Onofri) Oxford: OUP.
- [5] Boghossian, P. & Velleman, D. (1989). Color as a secondary quality. *Mind* 98 (January):81-103.
- [6] Boyle, Robert. (1664). *Experiments and Considerations Touching Colours*. London.
- [7] Buroker, J. (1991). Descartes on Sensible Qualities. *Journal of the History of Philosophy* 29 (4): 585-611.
- [8] Byrne, A. & Hilbert, D.R. (2003). Color realism and color science. *Behavioral and Brain Sciences* 26 (1):3-21.
- [9] Cassam, Q. (2010). Judging, believing and thinking. *Philosophical Issues* 20 (1): 80-95.
- [10] Cohen, J. (2009). *The Red and the Real: An Essay on Color Ontology*. Oxford University Press.
- [11] Chamberlain, C. (2019a). Color in a Material World: Margaret Cavendish against the Early Modern Mechanists. *Philosophical Review* 128 (3): 293-336.
- [12] ———(2019b). The body I call ‘mine’: A sense of bodily ownership in Descartes. *European Journal of Philosophy* 27 (1): 3-24.
- [13] Chisholm, R. M., & Feehan, T. D. (1977) The intent to deceive. *Journal of Philosophy* 74: 143-159.
- [14] Cottingham, J. (1990). XIII Descartes on Colour. *Proceedings of the Aristotelian Society* 90 (1): 231-246.

- [15] Descartes, R. (1996). *Oeuvres de Descartes*. Edited by Charles Adam and Paul Tannery. Paris: J. Vrin.
- [16] ———(1985). *The Philosophical Writings of Descartes* (Vols I and II). Translated by John Cottingham, Robert Stoothoff, and Dugald Murdoch. CUP.
- [17] ———(1991). *The Philosophical Writings of Descartes: The Correspondence* (Vol III). Translated by John Cottingham, Robert Stoothoff, Dugald Murdoch, and Anthony Kenny. Cambridge University Press.
- [18] De Rosa, R. (2007). A teleological account of cartesian sensations? *Synthese* 156 (2):311-336.
- [19] Downing, L (2011). Sensible Qualities and Material Bodies in Descartes and Boyle. In *Primary and Secondary Qualities: The Historical and Ongoing Debate*, edited by Lawrence Nolan. Oxford: OUP.
- [20] Garber, D. (1992). *Descartes's Metaphysical Physics*. University of Chicago Press.
- [21] Gottlieb, J., & Parvizian, S. (2018). Cartesian Imperativism. *Pacific Philosophical Quarterly* 99 (4): 702-725.
- [22] Gow, L. (2014). Colour. *Philosophy Compass* 9 (11):803-813.
- [23] Hardin, C. L. (1988). *Color for Philosophers*. Hackett.
- [24] Hatfield, G. (1986). The Senses and the Fleshless Eye: The Meditations as Cognitive Exercises. In Amelie Rorty (ed.), *Essays on Descartes' Meditations*. University of California Press. pp. 45-76.
- [25] ———(2013). Descartes on Sensory Representation, Objective Reality, and Material Falsity. In *Descartes' Meditations: A Critical Guide*, ed. Karen Detlefsen. Cambridge: CUP., pp. 127-50.
- [26] Johnston, M. (1992). How to speak of the colors. *Philosophical Studies* 68 (3):221-263.
- [27] MacKenzie, A (1990). Descartes on Sensory Representation. *Canadian Journal of Philosophy* 20 (Supplement) pp. 109–147.
- [28] Mahon, J, (2015). “The Definition of Lying and Deception”, *The Stanford Encyclopedia of Philosophy*, Edward N. Zalta (ed.).

- [29] Mandelbaum, E. (2016). Attitude, Inference, Association: On the Propositional Structure of Implicit Bias. *Nouûs* 50 (3): 629-658.
- [30] Maund, B. (2011). Colour eliminativism. In Lawrence Nolan (ed.), *Primary and Secondary Qualities: The Historical and Ongoing Debate*. Oxford: OUP.
- [31] Mendelovici, A. (2013). Reliable Misrepresentation and Tracking Theories of Mental Representation. *Philosophical Studies* 165 (2): 421- 443.
- [32] Nelson, A. (1997). Descartes’s ontology of thought. *Topoi* 16 (2): 163-178.
- [33] Nelson, A. & Cunning, D. (1999). Cognition and Modality in Descartes. *Acta Philosophica Fennica* 64: 137-154.
- [34] Nolan, L. (1998). Descartes’ Theory of Universals. *Philosophical Studies* 89 (2-3):161-180.
- [35] Nolan, L. (2011). Descartes on “What we call color”. In *Primary and Secondary Qualities: The Historical and Ongoing Debate*. OUP.
- [36] Parvizian, S. (2021). *Scientia*, diachronic certainty, and virtue. *Synthese* 198: 9165–9192
- [37] Patterson, S. (2016). Descartes on the Errors of the Senses. *Royal Institute of Philosophy Supplement* 78:73-108.
- [38] Perler, D. (2010). Does God deceive us? Skeptical hypotheses in late medieval epistemology. In Henrik Lagerlund (ed.), *Rethinking the History of Skepticism: The Missing Medieval Background*. Brill. pp. 103–171.
- [39] Quilty-Dunn, J. (2015). Believing in Perceiving: Known Illusions and the Classical Dual Component Theory. *Pacific Philosophical Quarterly* 96 (4): 550-575.
- [40] Ragland, C.P. (2006). Descartes on the Principle of Alternative Possibilities. *Journal of the History of Philosophy* 44 (3): 377–94.
- [41] Robinson, T. S. (2013). Descartes’s skeptical theism. *Religious Studies* 49 (4):515-527.
- [42] Schmaltz, T.M. (1991). Platonism and Descartes’ View of Immutable Essences. *Archiv fur Geschichte der Philosophie* 73, 129–170.
- [43] Schmaltz, T.M. (1995). Malebranche’s Cartesianism and Lockean Colors. *History of Philosophy Quarterly* 12 (4): 387-403.

- [44] Simmons, A. (2003). Descartes on the cognitive structure of sensory experience. *Philosophy and Phenomenological Research* 67 (3): 549-579.
- [45] ———(1999). Are Cartesian Sensations Representational? *Noûs* 33 (3): 347-369.
- [46] Stuart, J. (1986). Descartes' Proof of the External World. *History of Philosophy Quarterly* 3 (1): 19 - 28.