Fine-Tuning Divine Indifference

Chris Dorst and Kevin Dorst

July 2020

Draft—comments welcome!

Abstract

Given the laws of our universe, the initial conditions and cosmological constants had to be "fine-tuned" to result in life. Is this evidence for design? We argue that we should be uncertain whether an ideal agent would take it to be so—but that given such uncertainty, *we* should react to fine-tuning by boosting our confidence in design. The degree to which we should do so depends on our credences in controversial metaphysical issues.

Some old news: life exists.

Some new news: the old news was surprising—it turns out the laws of our universe are *stringent*, in the sense that the initial conditions and cosmological constants had to be "fine-tuned" to result in life.

The *fine-tuning argument* claims that the new news is evidence for a designer: when we learn that the laws are stringent—even already knowing that life exists—this provides further evidence for design.

We think this is right—but for subtle reasons. Weisberg (2010, 2012) offers a seemingly compelling argument that it's not, while White (2011) offers a seemingly compelling argument that it is. We'll show how their arguments are motivated by two different indifference assumptions, each of which has strong considerations in its favor. As a result, we should be unsure which assumption is correct. But here's the catch: when we should be uncertain which is correct, it follows that we *should* take stringency to be (further) evidence for design. How much we should do so depends on how much credence we should lend White's indifference assumption—which in turn is tied to metaphysical questions about nature of the physical laws and of the designer.

Alright, buckle up.

Let P be the (initial) credences of an ideally rational agent, D be the claim that there's a Designer, L be the claim that Life exists, and S be the claim that the laws are Stringent. The question White and Weisberg are trying to answer is whether, given life, an ideal agent would take stringency to be evidence for design: whether $P(D|LS) > P(D|L)$.

Both White and Weisberg are happy to accept the following two premises (formulated in Weisberg 2012):

Divine Intent: $P(L|D) = 1$.
Given a designer, it's certain that there'll be life.

Blind Indifference: $P(\cdot|L\overline{D})$ is uniform over life-worlds.
Given that life exists and there's no designer, each life-world should be treated as equally likely.

We'll accept them too. We think Divine Intent comes for free, given the sort of designer we have in mind. Meanwhile, Blind Indifference is motivated by the thought that absent a designer, it's a matter of "blind chance" which world comes about—thus $P(\cdot|\overline{D})$ is plausibly uniform over all worlds, and so $P(\cdot|L\overline{D})$ is uniform over life-worlds (Weisberg, 2012). (We'll loosen this assumption below.)

What White and Weisberg disagree on is what an ideal agent would be indifferent over given only the information that there is a designer: Weisberg thinks they would be indifferent over life-worlds, while White thinks they would be indifferent over levels of stringency.

To see the intuition behind Weisberg's argument, imagine the space of possible worlds arranged on a continuum, with the stringency of the laws increasing toward the left (Weisberg, 2012):

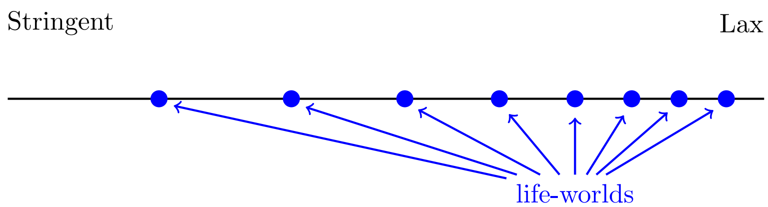

Figure 1: The Stringency Continuum

The dots on the line represent life-worlds. They become more common toward the right because as the laws become less stringent (more lax), there are more settings of the constants and initial conditions that will result in life.

Now our question: given (only) that there's a designer, how would an ideal agent distribute their credences? (What is $P(\cdot|D)$?) The picture suggests a natural answer: they'd treat each life-world as equally likely. After all, the only thing that's known about the designer is that she'll create life—so although she may well favor some life-worlds over others, there seems no basis for guessing how this favoring might go. This motivates Weisberg's crucial premise:

Divine *L*-Indifference: $P(\cdot|D)$ is uniform over life-worlds.
Given a designer, each life-world should be treated as equally likely.

It follows from these three premises that an ideal agent would *not* treat stringency as any further evidence of design: $P(D|LS) = P(D|L)$. After all, since both Blind Indifference and Divine L-Indifference treat each life-world as equally likely, learning that we're in a particular class of life-worlds (the stringent ones) doesn't tell in favor of either design or no-design.[1] Moreover, it also follows that given a designer, an ideal agent would think lax laws are more likely than stringent laws—for since there are more life-worlds to the right of the line, they would think a designer is more likely to wind up picking one of those worlds.

Does this mean that White was simply wrong, and that given a designer, an ideal agent would clearly *not* be indifferent over how stringent the laws are?

We don't think so: there's a different way of looking at things that motivates White's premise. Notice that there's something misleading about the above picture. At any given level of stringency, there are many different worlds—some of which contain life, others of which do not. Stringency measures the *proportion* of worlds with a given set of laws (but perhaps differing initial conditions or fundamental constants) that contain life (Weisberg, 2012).

So a more perspicuous diagram would be not a one-dimensional stringency continuum, but a two-dimensional stringency *space*, where the horizontal axis again arranges worlds according to stringency, and the vertical axis represents different settings of their fundamental constants and initial conditions (Figure 2). Every point in this diagram represents a world, and any vertical line represents an equivalence class of worlds at equal stringency. Intuitively, worlds with the same x-value have equally stringent laws, but different constants or initial conditions so that some of them are life-worlds, and others are not. Note that the proportion of life-worlds increases as the laws get less stringent.[2]

This picture highlights something hidden in the previous one: whatever level of stringency the designer picks, she can adjust ("fine-tune") the initial conditions and constants to make it a life-world.

Now return to our crucial question: given (only) that there's a designer, how would an ideal agent distribute their credences? (What is $P(\cdot|D)$?) This picture suggests that they might well treat each *stringency-class* as equally likely, knowing that the designer would simply adjust the initial conditions and constants to guarantee the emergence of life, regardless. This is especially intuitive if our rational agent imagines the designer to be primarily concerned with the fundamental structure of the universe, knowing that she can "fill in the details" as she sees fit.

Consider an analogy: a novelist is intent on writing a story in which—amongst many other things—a hero overcomes a challenge. Knowing only this, should we think it substantially more likely that the challenge will be easy to overcome? Plausibly not—since the novelist has complete control over the story, the difficulty of the challenge is a non-issue when it comes to ensuring that the hero *does* overcome it; thus the difficulty of the challenge will be fully determined by other goals the novelist has, and we have no

[1] More precisely: $P(D|LS) = P(D|L)$ iff $P(S|LD) = P(S|L)$, iff $P(S|LD) = P(S|L\overline{D})$. But Divine Intent implies that $P(\cdot|LD) = P(\cdot|D)$, and Divine L-Indifference says that this latter distribution is the same uniform distribution as Blind Indifference says $P(\cdot|L\overline{D})$ is. Thus $P(S|LD) = P(S|L\overline{D})$.

[2] It's not important to our argument that the proportion of life-worlds varies linearly with stringency, nor that it approaches anywhere near 100%. All that matters is that this proportion is monotonically increasing with laxity—though the precise rule given in footnote 4 will change if it increases non-linearly, as will the precise values calculated for $\lambda(S)$ and $\sigma(S)$ later.

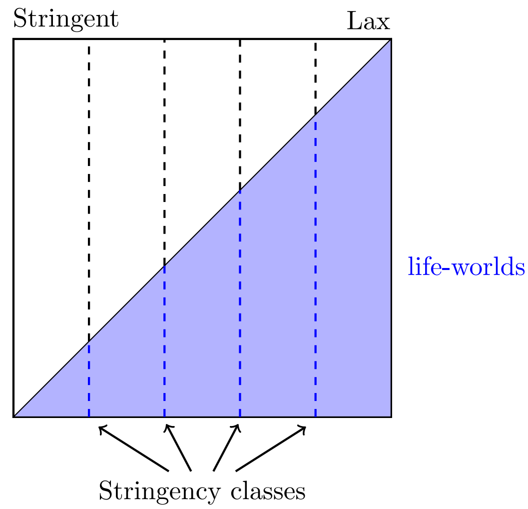

Figure 2: The Stringency Space

idea what those are. Likewise: since the designer has complete control over the world, a set of laws which makes life difficult to create is a non-issue when it comes to ensuring that life *is* created; thus the difficulty of creating life will be fully determined by other goals the designer has, and we have no idea what those are.

This, we think, is a perfectly intuitive picture—one that captures the "fine-tuning" intuition behind the original argument. It motivates White's premise[3]:

> **Divine *S*-Indifference:** $P(\cdot|D)$ is uniform over stringency-classes.
> Given a designer, each degree of stringency should be treated as equally likely.

It follows from Divine Intent, Blind Indifference, and Divine S-Indifference that an ideal agent *would* treat stringency as further evidence of design: $P(D|LS) > P(D|L)$.

To see why, consider Figure 3 (page 5). Given life-and-no-design, the rational credence in various levels of stringency increases with laxity (green line), since there are more life-worlds in the lax regions. Likewise given design-and-Divine-L-Indifference—this is why Weisberg's argument works, since those assumptions induce the same distribution as life-and-no-design. However, given design-and-Divine-S-Indifference, the rational credence is *uniform* over stringency levels (blue line). Therefore our credence that we end up in a stringent world given design and Divine S-Indifference ought to be higher than given life-and-no-design (the blue line is higher in the stringent region than the green one)—and hence if we accept Divine S-Indifference, stringency is (further) evidence for design: $P(S|LD) > P(S|L\overline{D})$, and hence $P(D|LS) > P(D|L)$.[4]

[3] We're deviating from White's (2011) formulation in ways that we think are natural (and that White would accept) given Weisberg's (2012) illuminating discussion.

[4] More precisely: let S be the proposition that we're in one of a set of stringency-classes centered on the left half of Figure 2. (It doesn't matter whether or not this region extends all the way to the left of the diagram.) Given Blind Indifference, $P(S|L\overline{D})$ will equal the proportion of the *area* of the life-

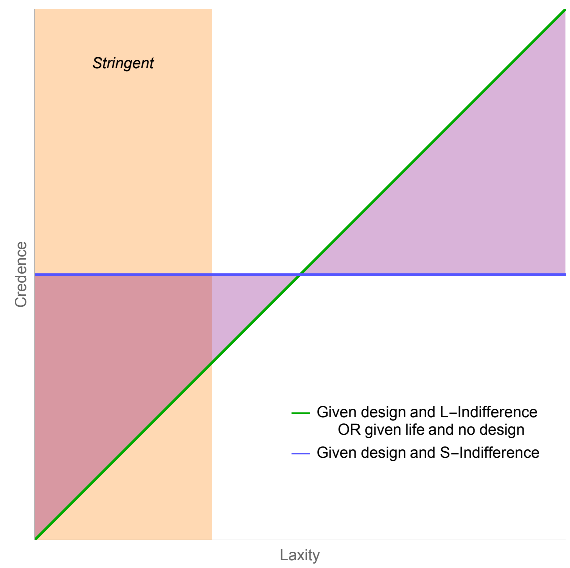

Figure 3: Credence in stringency-levels, given various hypotheses

So we have two plausible—but incompatible—hypotheses for how an ideal agent would distribute their credence given that there's a designer. We think it is pretty clear that non-ideal agents like us should be unsure which, if either, is correct (compare Roberts, 2012, p. 300). What should we do, given that uncertainty?

We should do what Bayesians always do: divide our credence amongst the various open hypotheses, consider how likely the possibilities are conditional on each hypothesis, and determine our overall opinion by averaging these conditional opinions. So, let C be a reasonable (but not necessarily ideal) initial credence function. We introduce it because we want to think about what people like *us*—who are uncertain what the ideal initial credence function P is—should think about the fine-tuning argument. We assume that C will defer to hypotheses about P.[5]

We want to know whether stringency should boost our confidence in design, given life—whether $C(D|LS) > C(D|L)$. This depends on whether we think that stringency is more likely given life-and-design than it is given life-and-no-design—whether $C(S|LD) > C(S|L\overline{D})$.

What should our credence be in stringency given life-and-no-design? (What is $C(S|L\overline{D})$?) Since we defer to P, this is determined by our credence in the various hypotheses about the ideal agent's distribution given life-and-no-design, i.e. $P(\cdot|L\overline{D})$. We take it that we can be sure that Blind Indifference is true—that given life-and-no-

triangle that SL takes up. Meanwhile, given Divine S-Indifference, $P(S|LD) = P(S|D)$ will equal the proportion of the *width* of the life-triangle that SL spans. Whenever S is centered on the smaller half of the life-triangle, the latter proportion will be larger than the former proportion, so $P(S|LD) > P(S|L\overline{D})$, hence $P(D|LS) > P(D|L)$.

[5]Precisely, where 'δ' is a rigid designator for a given probability function: $C(p|P = \delta) = \delta(p)$. It follows (by total probability) that $C(p|q \wedge P(p|q) = t) = t$.

design, an ideal agent would be uniform over life-worlds. Thus by deference, we should be too: $C(\cdot|L\overline{D})$ should be uniform over life-worlds.

What should our credence be in stringency given life-and-design? (What is $C(\cdot|LD)$?) By deference, this is again determined by our credence in the various hypotheses about the rational agent's credence distribution given life-and-*design*. We can put those hypotheses in three boxes: 'Divine L-Indifference' ('I_L'), 'Divine S-Indifference' ('I_S'), and 'something Else' ('E'). By total probability, $C(\cdot|LD)$ will be an average of its opinions conditional on each of these three hypotheses:

$$C(S|LD) \;=\; C\big(I_L|LD\big)\cdot C(S|LDI_L) \;+\; C\big(I_S|LD\big)\cdot C(S|LDI_S) \;+\; C\big(E|LD\big)\cdot C(S|LDE)$$

Letting λ and σ be the distributions posited by Divine L- and S-Indifference, respectively (such that, by Divine Intent, $\lambda(L) = \sigma(L) = 1$), deference then implies that your credence is an average of λ, σ, and your credence given that neither of them is ideal:

Divine Weighting:
$$C(S|LD) \;=\; C\big(I_L|LD\big)\cdot\lambda(S) \;+\; C\big(I_S|LD\big)\cdot\sigma(S) \;+\; C\big(E|LD\big)\cdot C(S|LDE)$$

Given reasonable assignments in Divine Weighting, it follows that $C(S|LD) > C(S|L\overline{D})$.

To see this, suppose for a moment that we're certain that one of White or Weisberg is right—either I_L or I_S is true. Then the third term in Divine Weighting drops out, and $C(S|LD)$ is a weighted average of $\lambda(S)$ and $\sigma(S)$. Since we can't be sure who is right, this average will assign positive weight to both terms. Graphically, this means that the likelihood of various levels of stringency is an average of the blue and green lines in Figure 3 (page 5)—i.e. a line passing through the purple region. Any such line assigns higher credence to the Stringent region than the green life-and-no-design line does, and therefore makes stringency more likely given design than given no design.[6]

Of course, we shouldn't be *certain* that one of White or Weisberg is right, meaning we can't pretend the $C(S|LDE)$ term isn't there. And its being there might be a problem: it could end up *lowering* the overall average given by Divine Weighting, since the denial of both L- and S-indifference (i.e. E) might make stringency even less likely. But as it turns out, this doesn't matter. Given *how* stringent we've learned the laws are, so long as we accord a tiny amount of credence to Divine S-Indifference, it follows that stringency is more to be expected given life-and-design than given life-and-no-design, *regardless* of the value of $C(S|LDE)$ (cf. Hawthorne and Isaacs, 2018).

An extremely generous estimate is that our universe is one in which around 1 in 10^{53} random settings of the parameters would result in life—we are scrunched down at the *very* bottom left of Figure 2 (Collins, 2003). So suppose that S is the claim that we are in a stringency-region of width $\epsilon > 0$, centered on a point where the green line has height $\frac{1}{10^{53}}$. Then what we need to evaluate is how the probability of S compares on both S-Indifference and L-Indifference—i.e. we need to evaluate $\sigma(S)$ and $\lambda(S)$. Since σ treats each vertical line in Figure 2 as equally likely, $\sigma(S)$ is the *width* of SL divided by

[6] More formally, we know from our exposition of White's argument that $\sigma(S) > \lambda(S)$, and we know that $C(S|L\overline{D}) = \lambda(S)$; therefore if $C(S|LD)$ is an average of $\sigma(S)$ and $\lambda(S)$, it is higher than $\lambda(S) = C(S|L\overline{D})$.

the *width* of L: $\epsilon/1 = \epsilon$. Meanwhile, since λ treats each point in the L-region of Figure 2 as equally likely, $\lambda(S)$ is the *area* of SL divided by the *area* of L. The area of L is $\frac{1}{2}$, and the area of SL is simply equal to the height of the life-region at the center of S (i.e. $\frac{1}{10^{53}}$) multiplied by the width of S—that is, $\lambda(S) = \frac{1/10^{53}}{1/2} \cdot \epsilon = \frac{2}{10^{53}} \cdot \epsilon$.[7]

So although the probability of S is small given either S-Indifference or L-Indifference, it is more than 50 orders of magnitude smaller given the latter: $\sigma(S) = \epsilon$ and $\lambda(S) = \frac{2}{10^{53}} \cdot \epsilon$. Moreover, recall that $\lambda(S)$ equals how confident you should be in S absent a designer: $\lambda(S) = C(S|L\overline{D})$. You should therefore think that, given life-and-no-design, it is monumentally unlikely that the universe would be this stringent. The question is whether it is any more likely given design.

To answer that, return to Divine Weighting. Since all of the terms are non-negative, any one of them—and in particular the S-Indifference one—gives us a lower-bound on $C(S|LD)$. It follows that so long as you have credence greater than $\frac{2}{10^{53}}$ in Divine S-Indifference given life-and-design (i.e. if $C(I_S|LD) > \frac{2}{10^{53}}$), you should (given life) take stringency to be more probable given design than not:

$$C(S|LD) \;\geq\; C(I_S|LD) \cdot \sigma(S) \;>\; \frac{2}{10^{53}} \cdot \epsilon \;=\; \lambda(S) \;=\; C(S|L\overline{D}).$$

Upshot: given how stringent we've learned the laws are, assigning even a minuscule credence to Divine S-Indifference suffices for the fine-tuning argument to work, regardless of what other hypotheses about the ideal prior distribution $P(\cdot|D)$ you leave open. The reason this works, again, is that while the S-Indifferent probability of S is low, the L-Indifferent probability is *far* lower. And since the latter equals your credence in S given life-and-no-design, it does not take much credence in Divine S-Indifference to push the probability assigned by Divine Weighting higher.

Summing up: to block the fine-tuning argument, it's not enough to merely think that Weisberg is right and White is wrong. Instead, you must think that White is so obviously wrong that you effectively assign credence *0* to him being right—for example, you must prefer to bet that this fair coin will land heads 175 times in a row than that White is right (since $1/2^{175} > 2/10^{53}$).

Balking at these odds, you might raise another worry about our argument. We've assumed Blind Indifference—that given life-and-no-design, an ideal agent would think each life-world is equally likely. But once we see that there are different versions of Divine Indifference, might there also be different versions of *Blind* Indifference? And might those differences matter?

There are, but they don't. Translating Divine L- and S-Indifference to the no-designer case would yield these two principles:

Blind *L*-Indifference: $P(\cdot|\overline{D})$ is uniform over life-worlds.

Blind *S*-Indifference: $P(\cdot|\overline{D})$ is uniform over stringency-classes.

[7] Similar calculations would go through varying the structure of the diagram—the point is simply that when you are in a very stringent region, the proportional width of the SL-region is *much* larger than its proportional area.

But *both* of these principles are plausibly true. For, as we've said, a natural hypothesis is that $P(\cdot|\overline{D})$ should be uniform over the entire stringency space (Figure 2): given only the information that there's no designer, an ideal agent wouldn't think that any level of stringency or life-(un)friendly settings of the conditions are more or less likely than any other.[8] That implies each of Blind L-, S-, and unadorned Indifference (page 2), and therefore that $C(\cdot|L\overline{D}) = \lambda(\cdot)$; our above argument goes through. In other words, it is only in the presence of Divine Intent that Divine L- and S-Indifference become incompatible. Since no parallel premise is plausible given no design (i.e. since $P(L|\overline{D}) < 1$), a parallel distinction amongst types of Blind Indifference makes no difference. So we don't think this is the place to object.

However, looking closer at these indifference distributions might raise another worry. Each of the three arguments (White's, Weisberg's, and ours) makes use of very demanding indifference assumptions: Divine S-Indifference uses exact uniformity over stringency classes, and both Divine L-Indifference and Blind Indifference use exact uniformity over life-worlds. You might think this exactness provides reason for skepticism about all of them: perhaps even if the basic idea of one of these principles is right, other considerations would cause the ideal agent to deviate somewhat from exact uniformity.

This seems fair enough. Interestingly, however, the different arguments have very different sensitivities to relaxation of the uniformity assumptions. Suppose that we change all of these "exact" uniformities to "rough" uniformities, so that Divine S-Indifference says the ideal distribution is *roughly* uniform over stringency classes, and both Blind Indifference and Divine L-Indifference say that it's *roughly* uniform over life-worlds.

White's argument and our own are insensitive to this change. This is because—again, given how stringent we've learned the laws are—Divine S-Indifference makes stringency *so* much more likely than Blind Indifference that even significant deviations from each would render the comparison between the two unchanged. (In the left-most part of Figure 3, large fluctuations in the blue and green lines would still leave the former much higher than the latter.) For White's argument, all that matters is that that the value of $\sigma(S)$ is higher than the value of $\lambda(S)$. For us, so long as the value of $\sigma(S)$ is *substantially* higher than the value of $\lambda(S)$—as surely it still will be even on these relaxed assumptions—we only need to accord the former a minimal amount of weight in Divine Weighting to ensure that $C(S|LD) > C(S|L\overline{D})$.

Weisberg's argument, in contrast, relies on an *exact* match between the indifference posited by Blind Indifference and that posited by L-Indifference. If the (rough) Divine L-Indifference distribution ends up according even slightly higher probability to stringency than the (rough) Blind Indifference distribution does, the conclusion that stringency is irrelevant to design no longer goes through. In contrast to this fragility, then, the robustness of our argument strikes us as an advantage.

Summing up: for non-ideal agents like us—who should be unsure whether Divine

[8] More carefully: each equally-sized region of stringency space is assigned the same probability. This is compatible with thinking some types of worlds are a priori more probable than others; it just means that we don't have reason to suspect that such worlds are unevenly spread throughout stringency space.

S-Indifference is true—the fine-tuning argument succeeds: upon learning the laws are stringent, we should (further) boost our confidence in design.

Okay—but *how much*?

The answer, we think, hinges on how confident we should be in a variety of controversial metaphysical theses about the nature of the designer and the laws. Theses that make Divine S-Indifference more plausible increase its weight in Divine Weighting, thereby further boosting the likelihood of stringency given life-and-design above its likelihood given life-and-no-design.

For example, a theistic picture of a designer—who not only creates the universe, but also intervenes to make sure things go according to plan—would appear to motivate Divine S-Indifference. For on that picture, no matter what laws the universe has and no matter how it started, the designer would tweak the particular facts to ensure life ends up arising. In contrast, a deistic picture of a designer—who simply creates the universe and steps back—lends support to Divine L-Indifference (and so takes it away from S-Indifference). For on that picture, the designer picks a world based on whether, unaltered, it'll play out to contain life.

Similarly, on a governing conception of laws, the laws are ontologically prior to the rest of the goings-on in the universe. On such a picture, it seems quite reasonable that a designer would first select the laws—the fundamentals of the world—and then fine-tune the initial conditions and constants to permit life; thus the governing conception supports Divine S-Indifference. By contrast, Humeanism about laws would seem to tell against Divine S-Indifference, since on this view the laws are merely patterns in the particular facts that obtain throughout the universe; they are not a separate variable that can be manipulated before determining how the rest of the universe will go.

These metaphysical debates bear on the fine-tuning argument by motivating different values for the $C(I_S|LD)$ term in Divine Weighting; the larger it is, the more significant the discovery of stringency for the existence of a designer. No doubt other debates will bear on this weighting as well.

Where does this leave us? Since none of us should be sure of what an ideal agent would think given only the existence of a designer, we all should take stringency to provide some (further) evidence of design—but the degree to which we should do so depends on our opinions about further metaphysical debates.

Upshot: the fine-tuning argument succeeds. But the *degree* to which it succeeds is a subtle, interesting, and open question.

References

Collins, Robin, 2003. 'Evidence for fine-tuning: Robin Collins'. In Neil A Manson, ed., *God and Design*, 193–214. Routledge.

Hawthorne, John and Isaacs, Yoaav, 2018. 'Fine-tuning fine-tuning'. *Knowledge, Belief, and God: New Insights in Religious Epistemology*, 136–168.

Roberts, John T., 2012. 'Fine-tuning and the infrared bull's-eye'. *Philosophical Studies*, 160(2):287–303.

Weisberg, Jonathan, 2010. 'A note on design: What's fine-tuning got to do with it?' *Analysis*, 70(3):431–438.

———, 2012. 'The argument from divine indifference'. *Analysis*, 72(4):707–714.

White, Roger, 2011. 'What fine-tuning's got to do with it: A reply to Weisberg'. *Analysis*, 71(4):676–679.