

# *Philosophizing out of bounds*

**Jennifer Nado**

**Philosophical Studies**

An International Journal for Philosophy  
in the Analytic Tradition

ISSN 0031-8116

Philos Stud

DOI 10.1007/s11098-020-01582-0



**Your article is protected by copyright and all rights are held exclusively by Springer Nature B.V.. This e-offprint is for personal use only and shall not be self-archived in electronic repositories. If you wish to self-archive your article, please use the accepted manuscript version for posting on your own website. You may further deposit the accepted manuscript version in any repository, provided it is only made publicly available 12 months after official publication or later and provided acknowledgement is given to the original source of publication and a link is inserted to the published article on Springer's website. The link must be accompanied by the following text: "The final publication is available at [link.springer.com](http://link.springer.com)".**



# Philosophizing out of bounds

Jennifer Nado<sup>1</sup>

Accepted: 12 October 2020  
© Springer Nature B.V. 2020

## 1 Introduction

The experimental philosophy studies that began to emerge in the first years of the twenty-first century were catalysts to a lot of methodological anxiety. Participants' responses to thought experiments appeared to be sensitive to all sorts of funny things, such as cultural background or socioeconomic status, or how a vignette was framed. Such studies seemed to compel experimentalists towards a dramatic, pessimistic conclusion: the intuitions on which philosophical method rests are too unreliable to serve as evidence for philosophical conclusions. The armchair must burn.

It's been about two decades since then, and the so-called 'negative program' of experimental philosophy has gotten a lot more sophisticated. There are more studies now, of course—one only needs to look at Machery's incredibly useful literature survey to see that—but there's also been much more attention paid to exactly how the experimentalists' arguments should be construed. Machery's book is an exemplary contribution to this ongoing project of sharpening the experimentalist critique.

One criticism of the initial, 'naïve' formulation of the experimentalist argument which I suspect looms large in Machery's mind is due to Timothy Williamson.<sup>1</sup> Williamson has pressed against the notion of 'intuition' that undergirds naïve experimentalism, noting that the cognitive capacities that underlie 'intuitive' judgment are not particular to philosophy. They simply reflect, for instance, an

---

<sup>1</sup> See Williamson (2007, 2015).

---

✉ Jennifer Nado  
nado@hku.hk

<sup>1</sup> University of Hong Kong, Pok Fu Lam, Hong Kong

ability to apply concepts—and are thus plausibly shared with much of the bulk of everyday cognition. This overlap between philosophical and quotidian judgment has dire consequences for any attempt to reject use of ‘intuition’ wholesale—by such lights, even judgments like ‘I know where I left my keys’ ought to be deemed *verboten*.

There have been a handful of attempts to prevent the sceptical collapse that Williamson takes experimentalism to invite. One early suggestion, due to Weinberg (2007), was to move away from unreliability as the basis of anti-intuition critique to some feature which ordinary judgment purportedly lacks (Weinberg’s proposed substitute was ‘hopelessness’). A later proposal, offered by both Alexander and Weinberg (2014) and by myself (Nado 2015), was to note that the activity of philosophical theory-building is simply much more error-sensitive than our ordinary, day-to-day cognitive tasks. More broadly, there has been a marked tendency in experimentalist writing to drop the ‘intuition’ category in favour of less loaded terms such as ‘philosophical judgment’—though doubtless this latter has also been fuelled by recent arguments by Cappelen (2012) and Deutsch (2015) (among others) that deny that intuition plays a significant role in philosophical method.

Machery’s own revamp of naïve experimentalism likewise eschews the ‘intuition’ label, preferring to simply speak of judgments elicited by philosophical cases. Machery sticks with unreliability as the relevant epistemic defect, and focuses on identifying features of the *cases* philosophers employ which he holds to magnify the unreliability of our judgments—such features are delightfully labelled the ‘disturbing characteristics’. In essence, Machery holds that philosophical theorizing involves using perfectly respectable cognitive capacities outside of their proper domains. If successful, Machery’s approach would quarantine philosophical activity away from more acceptable instances of ordinary judgment, stemming the potential sceptical epidemic that naïve experimentalism threatened to unleash.

The conclusions Machery draws from this approach,<sup>2</sup> however, are nearly as pessimistic as those of the naïve critique—indeed in some ways *more* pessimistic. The core of Machery’s position is self-confessed *modal scepticism*; since many modal philosophical claims are outside of our judgments’ proper domains, “we cannot know many of the metaphysical possibilities and necessities of philosophical interest” (Machery 2017, p. 2). As a result, philosophical issues requiring judgment on ‘modally immodest’ claims are “beyond our epistemic reach”, and Machery recommends that we “set them aside in order to turn our attention toward issues we can fruitfully theorize about” (Machery 2017, p. 245). Thus, though naïve experimentalists might rest content with critiquing one particular philosophical tool (‘intuition’), Machery submits that entire philosophical *projects* are ultimately ‘out of bounds’—including, he writes, debates over physicalism, free will, and the analysis of knowledge.

---

<sup>2</sup> I focus in the following on Machery’s ‘unreliability’ argument; he also uses arguments based on dogmatism and parochialism to support the same conclusions. However, most of my concerns about the former also apply to the latter.

It's not all doom and gloom, however. Machery offers consolation in the form of a modified, naturalistic version of conceptual analysis which he holds to be within philosophy's proper bounds. Experimentally informed investigation of concepts, construed as psychological entities, is still achievable. And, this empirical investigation of our psychology can sometimes be employed in service of the sort of prescriptive *modification* of concepts that recent literature dubs 'conceptual engineering'.

My own philosophical sympathies align not only with the negative experimentalist's empirically-based suspicion of 'traditional' methodology, but also with the embrace of conceptual engineering as an improvement on said methodology. As such, I find relatively little to disagree with in the overall picture Machery offers. But I am, ultimately, a bit more of an optimist—not only about philosophy as it currently stands, but about the extent of our potential epistemic reach. In what follows, I'll argue that while the foundations of Machery's critique are compelling, several of the more eyebrow-raising aspects of his conclusions are unwarranted.

## 2 Quarantining the Philosophical

Let's begin with points of agreement. Machery devotes substantial space to carefully arguing that extant experimental philosophy studies suffice to prompt genuine worry about at least some subset of our philosophical judgments. Concerns about effect size, replicability and the like largely wash out in the face of the accumulated experimental evidence. I agree. Second, I share with Williamson and with Machery the view that philosophical case-judgments employ cognitive capacities that are also used in ordinary applications of concepts in everyday judgment. Third, I'm on board with Machery's claim that unreliability is increased by certain aspects of the sorts of cases philosophers often rely on: the 'disturbing characteristics'.

One such 'disturbing' feature is unusualness, in the sense of being rarely encountered (as opposed to being atypical—chickens are atypical but frequently encountered birds). A second is 'pulling apart what usually goes together'—that is, creating a case in which property A is present but property B is not, even though A typically goes along with B. An example Machery gives is the footbridge case "pulling apart engaging in physical violence and doing more harm than good". The third feature Machery discusses is the 'entanglement' of essential features of a described case with superficial, inessential aspects of its narrative content—for instance, whether the victim of our utilitarian footbridge-toppling is named 'Tyrone Payton' or 'Chip Ellsworth III' (Uhlmann et al. 2009). Machery recruits evidence from psychology to back the claim that these characteristics decrease reliability in judgment, leading judgments on cases with such characteristics to be out of our cognitive capacities' proper bounds.

It's at this point where I begin to have concerns. Machery claims that disturbing characteristics increase the likelihood of demographic or presentational variation effects. Crucially, this is meant to undergird an inductive move from cases which experimentalists have studied, to judgments on philosophical cases as a whole—for,

Machery argues, philosophical cases are typically awash in these variation-provoking disturbing characteristics. We should, then, expect most philosophical case-judgments to display inappropriate variability, and thus substantial unreliability. As part of this inductive move, Machery argues that “as far as we have information, reliability is invariant under partitioning of the class of judgments elicited by philosophical cases” (Machery 2017, p. 99). In other words, philosophical case-judgment is the appropriate kind over which to generalize—for we have no information that would suggest, e.g., that there are significant differences in the reliability of epistemic vs moral judgment, or ‘snap’ vs reflective judgment.

This nicely isolates Machery’s intended target, but are alternate partitionings really so quickly dismissed? Machery’s characterization of philosophical case-judgment is ‘minimalist’—he rejects not only claims that such judgment essentially involves e.g. conceptual competence or distinctive phenomenology, but also characterizations in terms of cognitive characteristics like speed or lack of conscious inference. He dismisses the latter characterization largely because “experimental evidence suggests that the judgments elicited by at least some philosophical cases are consciously inferred” (Machery 2017, p. 42). Fair enough—but isolating the targeted set of judgments purely via characteristics of the cases judged, rather than characteristics of the cognitive processes used to judge them, results in a class of judgments that are undeniably extremely heterogeneous.

Suppose I come to have a view on a ‘disturbing’ philosophical case based on an inference from further theoretical commitments, or on some other chain of argumentation. Are all such judgments to be dismissed along with the more ‘intuitive’ snap judgments that x-phi originally aimed to target? For instance: Weatherson (2003) argues that we ought to count Gettier cases as known on grounds of the naturalness and simplicity of JTB. This is a judgment on a disturbing philosophical case. But it is wholly based on further judgments/commitments, none of which obviously hang on philosophical case-judgments. Machery might protest that in this example Weatherson is not employing the *method* of cases, in that he is not using the Gettier case-judgment as support for an analysis of knowledge (rather, the support is in the other direction). But the *use* of the judgment isn’t relevant here—Machery’s argument looks only at whether the judgment is unreliable. Machery’s argument states that we ought to “refrain from making a judgment of a particular kind K... when most judgments of this kind are [unreliable]” (Machery 2017, p. 102). He argues that philosophical case-judgment is one such kind. And that kind is one to which Weatherson’s judgment belongs.<sup>3</sup>

If Machery is inclined to permit at least some explicitly inference-based philosophical case-judgments, a partition must be made—but Machery’s minimalism gives us scant construction material. Minimalism in essence claims that philosophical judgments aren’t distinct from everyday judgments, but of course everyday judgments can be based on argumentation or consistency with prior commitments, too. On the other hand, if Machery does intend to target *all*

---

<sup>3</sup> Machery does say that judgments belonging to unreliable kinds may be used when said judgment is known to be an exception, but it’s unlikely that Weatherson’s controversial judgment meets that criterion.

judgments on disturbing cases, the case for differing reliability levels under partitioning looks pretty strong. Perhaps Machery is right that we do not currently have enough information to make pronouncements of the form ‘philosophical case-judgments with feature *x* are substantially more reliable than philosophical case-judgments considered as a whole’. But given the plausible heterogeneity of the cognitive processes leading to said case-judgments (and I have not even touched on the plausible heterogeneity *within* ‘snap’ or ‘unreflective’ judgment), we have *prima facie* reason to assume that *some* such cognitive process types will substantially differ in their vulnerability to the disturbing characteristics—and that, with enough empirical work, those cognitive process types may be eventually identified.

Whether any such types will be sufficiently reliable to permit their use is a second question; and again, one that rests on empirical information we arguably do not yet have. What are we to do in the meantime, in our current state of ignorance? Machery argues that the correct move is to abandon the whole lot. He motivates this by analogy—“if I know that most eggs in a pack are rotten, the reasonable thing in the absence of further information is to throw the *whole* pack” (Machery 2017, p. 93). Well, yes; but rotten eggs make one sick, and the only downside of tossing the pack is possibly missing breakfast. I’d toss the pack if only *one* was rotten. Machery is aware that cost/benefit analysis is relevant here, but claims that “making a false judgment in response to a philosophical case is costly: one develops erroneous philosophical theories” (Machery 2017, p. 93). Yet sickening as false philosophical theories might be to some, Machery’s alternative is arguably a greater cost: at best, a radical restriction of our philosophical toolbox, and at worst (if Machery’s right), a removal of all hope of settling long-standing philosophical questions. Is it really better to refrain from theorizing entirely rather than run a moderate risk of getting things wrong? Legions of weathermen would likely protest otherwise.

The costs of abandoning the ‘whole pack of eggs’ extend beyond the ivory tower, too. Philosophical cases aren’t the sole domain of philosophers. Ordinary folk talk of ethics, of freedom, of knowledge, of beauty, of mental states, of causation. And they often judge whether such properties apply to cases—sometimes ‘disturbing’ cases. What’s worse: we already know that at least some such ‘everyday’ judgments vary as a function of demographic and presentational features. To take the most obvious case, ethical judgments in non-philosophical contexts are undeniably affected by cultural, religious, and socio-economic background.

Practical matters often hang on such variation-sensitive judgments: for instance, consider debates over whether European countries ought to ban burqas, or whether pharmacists ought to be allowed to refuse to fill birth control prescriptions on religious grounds, or whether to permit euthanasia or the death penalty. Do the cases that ordinary folk are likely to consider as part of these debates display disturbing characteristics? Arguably they do, at least to some degree. It’s unusual, for instance, for a pharmacist to refuse to fill a prescription; and euthanasia pulls apart the normally co-occurring features of providing medical care and aiming to preserve life. As for entanglement with superficial characteristics, the actual-world nature of many of the relevant cases inevitably means that they will come pre-entwined with all sorts of irrelevant details—if implied race is enough to make footbridge

judgments risk substantial error, then burqa-ban debates are doomed to be risky. But it's a risk we must take; the sort of considered suspension of judgment that Machery recommends isn't really an option here.

And that's to stick with the philosophical. If the argument for rejecting philosophical case-judgments is grounded in the relevant cases' possession of disturbing characteristics, what of non-philosophical cases of concept application? For instance: the property of 'pulling apart what usually goes together' renders cases atypical, and Machery's case against said property is based on findings of reduced levels of consistency and consensus on atypical members of ordinary categories. As an example, Machery invokes the fact that people disagree over whether a lift is a vehicle—ought we, then, to refrain from judging such non-philosophical cases too? Judgments on unusual and/or atypical cases may make up a relatively small slice of ordinary cognitive activity, and surely we are indeed less reliable on such cases, but to *prohibit* such judgments seems fairly extreme.

There are various maneuverings that Machery might make to keep the philosophical and the everyday separate here—unusualness and atypicality come in degrees, after all, so perhaps we might draw a line between classifying a taco as a sandwich and evaluating midnight-movie sci-fi scenarios such as Swampman.<sup>4</sup> But even supposing we could draw a non-arbitrary line that would spare enough of everyday cognition, it would almost certainly spare a substantial amount of philosophical cases, too—just within ethics, the mob and the magistrate, trolley cases, organ-harvesting cases and the like aren't obviously more 'disturbing' than the burqa-ban and birth control cases discussed above. My complaint here, I should note, is self-consciously Williamsonian: it is that many of the judgments non-philosophers make in ordinary contexts possess at least some measure of the disturbing characteristics, and many philosophical case-judgments—even ones explicitly within Machery's sights—are no more disturbing than these.

### 3 The Perils of Modal Skepticism

Let's move now from the particular to the general. Machery doesn't merely claim that philosophical case-judgments are unreliable; he argues that a form of modal scepticism follows. Machery doesn't, however, claim that *every* modal claim is out of reach; he explicitly restricts his claim to "metaphysical modalities of philosophical interest" (Machery 2017, p. 187), thus permitting knowledge of nomological necessities and possibilities, and at least some non-philosophical metaphysical modal claims. Indeed, he notes that we may even have knowledge of a few philosophically interesting modal necessities—he offers as a potential example the claim that everything is necessarily identical to itself. But large swaths of

---

<sup>4</sup> Machery does make some moves in this vein—he writes that unusualness is only a problem when the accuracy of judgment depends on considering the unusual feature (thus making the process of adding two pink elephants to two pink elephants not unusual in the relevant sense), for instance. This move does not, however, defang the cases discussed above.



philosophically relevant modal facts are held to be beyond our reach; enough, Machery claims, to prevent theorizing on many traditional philosophical questions.

It's worth noting that the primary target here seems to be claims of metaphysical *necessity*—non-negated ones, at that. Presuming Machery's going to allow us the inference from 'P' to 'possibly P', and from 'possibly P' to 'not necessarily not P', we can claim for ourselves a pretty nice store of philosophically relevant metaphysical modal knowledge. Rather, the target is claims like 'necessarily, if S knows that P then S has a true belief that P', or 'necessarily, if a belief is both true and reliably formed then it is known'.

But what of Machery's own claim that "We cannot know many of the metaphysical possibilities and necessities of philosophical interest" (Machery 2017, p. 2)? On an uncharitable reading, this claim comes uncomfortably close to the self-undermining 'necessarily, beliefs about metaphysical possibilities and necessities of philosophical interest are not knowledge'. Perhaps Machery would hold that this necessity is not one of the aforementioned 'many'—but without further motivation this would be ad hoc, particularly because the analysis of knowledge is explicitly on Machery's hit-list.

Slightly more charitably, perhaps Machery simply means to claim that knowledge of most philosophical necessities/possibilities is *nomologically* impossible. But Machery's argument for modal scepticism is grounded by appeal to the unreliability of case-judgment; and thus *prima facie* on the claim that some degree of reliability is a necessary condition for knowledge. And it would strike me as deeply odd to hold that we are justified in believing that reliability is required for knowledge in all nomologically possible cases, but that that justification disappears when the laws of nature aren't held fixed. The laws of nature simply seem to have nothing to do with it. Our opposition to the commingling of unreliability and knowledge doesn't seem to be implied by, say, appeal to our best scientific theories; it seems to simply be rooted in our ordinary capacity for identifying knowledge. But if that's right, then Machery likely can't even help himself to the *nomological* necessity that's needed here. Machery's modal scepticism is fuelled by the unreliability of philosophical case-judgment. But the vast majority of philosophical cases, even disturbing ones, are nomologically possible. So why is it only knowledge of *metaphysical* possibility and necessity that gets the axe?

Finally, suppose we read Machery's claim as grounded in counterfactual knowledge: in all worlds reasonably close to the actual one, we fail to know (most) philosophical metaphysically-modal claims. But again we've got a similar issue—many philosophical cases are really rather 'nearby'. Gettier cases, trolley problems, and Gödel name-swaps don't exactly strain at the boundaries of nomological possibility. So again, if we are denying ourselves cases like those, even philosophically-relevant generalizations over counterfactuals may be hard to come by. It's worth noting too that, for Machery's argument to work, he doesn't need merely to rule out the co-instantiation of knowledge and 'at-chance' levels of reliability. The level of unreliability philosophical case-judgments display is plausibly moderate, perhaps even close to borderline acceptable. So here's the crucial question: does application of our ordinary knowledge-attribution abilities purely 'within bounds' suffice to justify the reliability requirement needed for

Machery's modal skepticism? If not, is there another route to the justification of this reliability requirement that will *not* also enable us to pursue the sorts of philosophical theorizing Machery condemns?

#### 4 Pushing our Boundaries

I've just made a fair amount of fuss over a point that I think is actually easily avoidable. Given that knowledge is one of these phenomena whose modal features Machery deems largely inaccessible to us, why frame the issues in terms of knowledge in the first place? After all, Machery has another tool at his disposal—naturalized, prescriptive conceptual analysis. What's critical, really, is not whether claims like 'knowledge requires being non-Gettiered' amount to knowledge; it's whether our epistemic status with regard to such claims is sufficiently poor that we should abandon some current aspects of our philosophical practice. Why not, then, simply prescribe a concept—knowledge\*—that reflects the level of epistemic quality a judgment needs to have in order to underwrite philosophical theorizing? Of course, if a judgment of the form 'knowledge\* requires reliability' itself requires philosophical case-judgment, we may still be in trouble; but Machery seems confident that prescriptive analysis can be pursued without reliance on cases. I'm less sanguine, but that's a kettle of fish I've no space to boil.

Even with a tailor-made knowledge\* concept, the stubborn overlap between everyday and philosophical judgment means that Machery's conclusions will likely still need some tinkering. One strategy, which I favour (see e.g. Nado 2017), would be to make the threshold for knowledge\* higher than that for knowledge, allowing that 'disturbing' case-judgments might be good enough for everyday activity while being insufficiently reliable for philosophical purposes. I'd also suggest unlinking the experimentalist critique from the otherwise plausible 'minimalist' characterization of philosophical judgment, and revisiting the category of quick, not-consciously-inferential 'snap' judgments. Most, but not all, experimental studies at least *seem* to involve such judgments. If things are indeed as they seem (and perhaps measures like response time might increase our confidence here), why not restrict the critique to these? Suppose we call these 'intuitions'. Then (a) many but not all ordinary judgments involve intuitions; (b) many but not all philosophical case-judgments involve intuitions. Limiting the critique to such intuitions would then allow us to give a pass, for instance, to arguments like those of Weatherson (2003).

Unfortunately, this is still not enough, since the category of intuition itself presumably subsumes multiple different cognitive processes whose reliability levels likely vary. A final needed adjustment, I think, is to step away from the claim that we ought to unilaterally refrain from a target class of judgments, in favour of a more graded conclusion: the more disturbing characteristics a case involves, the less confident we should be in our judgments in response to said case. We should, consequently, be more willing to 'bite the bullet', to prescribe counterintuitive modifications to our concepts, and so forth. Our level of confidence may also be influenced by other factors, including any information we have regarding the cognitive processes involved in producing the judgment and their relative

susceptibility to variation. At current such information is limited, but as our understanding of the processes underlying philosophical judgment increases, more sophisticated responses to the flaws of our cognitive capacities will become available. In the meantime, we can philosophize out of bounds—perhaps with just a little less bravado.

## References

- Alexander, J., & Weinberg, J. (2014). The “unreliability” of epistemic intuitions. In E. Machery (Ed.), *Current controversies in experimental philosophy* (pp. 128–145). New York: Routledge.
- Cappelen, H. (2012). *Philosophy without intuitions*. Oxford: Oxford University Press.
- Deutsch, M. (2015). *The myth of the intuitive*. Cambridge, MA: MIT Press.
- Machery, E. (2017). *Philosophy within its proper bounds*. New York: Oxford University Press.
- Nado, J. (2015). Intuition, philosophical theorising, and the threat of Scepticism. In E. Fischer & J. Collins (Eds.), *Experimental philosophy, rationalism, and naturalism: Rethinking philosophical method* (pp. 204–221). London: Routledge.
- Nado, J. (2017). Knowledge is not enough. *Australasian Journal of Philosophy*, 95(4), 658–672.
- Uhlmann, E., Pizarro, D., Tannenbaum, D., & Ditto, P. (2009). The motivated use of moral principles. *Judgment and Decision Making*, 4, 476–491.
- Weatherson, B. (2003). What good are counterexamples? *Philosophical Studies*, 115, 1–31.
- Weinberg, J. (2007). How to challenge intuitions empirically without risking skepticism. *Midwest Studies in Philosophy*, 31, 318–343.
- Williamson, T. (2007). *The philosophy of philosophy*. Oxford: Blackwell.
- Williamson, T. (2016). Philosophical criticisms of experimental philosophy. In J. Sytsma & W. Buckwalter (Eds.), *A companion to experimental philosophy* (pp. 22–35). Oxford: Wiley Blackwell.

**Publisher’s Note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.