

The Self-Effacement Gambit*

Penultimate version of a paper forthcoming in *Res Philosophica* - Cite Appropriately

Jack Woods

Philosophical arguments usually are and nearly always should be abductive. From Lewis's claim that philosophy is really a game of weighing costs,¹ to the rise of anti-exceptionalism in logic,² to the near universal reliance on methods like reflective equilibrium in ethics and metaethics,³ philosophers are starting to recognize that often the best we can do in theorizing some phenomena is put forward our best overall account of it, warts and all. This is especially true in esoteric areas like logic, aesthetics, mathematics, and morality where the data to be explained is often based in our stubborn intuitions.

While this methodological shift is welcome, it's not without problems. Abductive arguments involve significant theoretical resources which themselves can be part of what's being disputed. This means that we will sometimes find otherwise good arguments which suggest their own grounds are problematic. In particular, sometimes revising our beliefs on the basis of such an argument can undermine the very justification we used in that argument.

This feature, which I'll call *self-effacingness*, occurs most dramatically in arguments against our standing views on the esoteric subject matters mentioned above:

*Thanks to Tim Button, Justin Clarke-Doane, Catarina Dutilh-Novaes, Billy Dunaway, Edward Elliot, Daniel Elstein, Graeme Forbes, Marcus Giaquinto, Jessica Isserow, Simon Kirchin, Mary Leng, Jimmy Lenman, Barry Maguire, Erum Naqvi, Huw Price, Mark van Roojen, Josh Schechter, Paulina Sliwa, Robert Stern, Mike Titelbaum, Robbie Williams, Daniel Wodak, and Gözde Yıldırım for helpful discussion of earlier versions. Thanks also to audiences at the Epistemology of Mathematics: Knowledge, Proof, and Explanation meeting at Leeds, the Maths and Morals meeting at Kent, the Cambridge Moral Sciences Club, the Sheffield Undergraduate conference, and the St. Louis conference on Reasons and Rationality for useful discussion.

¹See the introduction to his (1983) for a classic statement.

²See (Hjortland 2017), (Priest 2006, 2016), (Russell 2015), and Williamson (2017, manuscript). There are still serious problems for implementing anti-exceptionalist credos about logic (Woods forthcoming-a, forthcoming-b).

³See (Scanlon 2014) for trenchant defense and (McPherson 2015) for worries.

logic, mathematics, aesthetics, and morality. This is because these subjects all play a role in how we reason abductively. This isn't an idle fact; we can resist some challenges to our standing beliefs about these subject matters exactly because the challenges are self-effacing. The self-effacing character of certain arguments is thus a benefit and limitation of the abductive turn and deserves serious attention. I aim to give it the attention it deserves.

I'll start by distinguishing two types of self-effacement, giving detailed examples of both. The first type occurs when accepting a conclusion undermines the materials we're abducting on; the second when accepting a conclusion undermines abductive methodology itself. Each of logic, mathematics, aesthetics, and morality allow self-effacement in at least one of these two ways. In fact, under plausible hypotheses, otherwise good self-effacing arguments occur in each.

These arguments take the form of challenges to our beliefs about these subjects *en bloc*. Recent examples include debunking arguments like those articulated by Harman (1977), Joyce (2001), and Street (2006), as well as more general skeptical challenges like those posed by Cohen (2000). I'll focus on these global challenges in what follows, leaving discussion of the less interesting case of local self-effacement largely to the side.

I'll clarify how, why, and most importantly when *self-effacement* blocks this kind of challenge to our standing beliefs. Unfortunately, we can still use many blocked challenges as *premises* of new challenges to our beliefs. These blocked challenges are still cases where our beliefs about ϕ say that they're not the best overall beliefs about ϕ . That's a significant vice for any set of beliefs. So we can sometimes use the viciousness of ϕ to argue that we ought to revise ϕ *without* self-effacement.

Our beliefs about morality and aesthetics only gain limited protection from challenges by means of self-effacement for this reason. This is especially plausible when we focus on the costs of wholesale revision or rejection of our moral and aesthetic beliefs to our abductive practice.⁴ The ability to use abduction survives wholesale revision or rejection of these subjects.

Mathematics and logic fare better. We use logic and mathematics not only in

⁴I'll put aside practical costs and theoretical costs unrelated to abductive methodology. It's worth noting that the costs of abandoning our standing moral or aesthetic beliefs might be different when we factor these in. But, since the costs to abduction of abandoning mathematics and logic are enormous, stable, and relatively uncontentious, I'll put these other costs aside for simplicity.

justifying the “inputs” to an abductive argument, but also in aggregating the facts about theoretical virtues so input. I argue that this fact shows that to reject our logical and mathematical beliefs *en bloc* would be tantamount to abandoning the abductive method entirely. So, while the cogency of any self-effacing argument is problematic, the extreme cost of coming to doubt our beliefs about mathematics and logic suffice to ignore the otherwise problematic existence of self-effacement.

This gives rise to a more nuanced “abductivized” blocking strategy. The overwhelming cost of rejecting our current mathematical and logical beliefs overwhelm the viciousness of self-effacing arguments. The costs of rejecting our current moral and aesthetic beliefs aren’t clearly high enough to do so. So, mathematics and logic are on firmer ground than morality, even if certain challenges to morality are also self-effacing.

In short, if we go abductive about the costs of abandoning aspects of our abductive methodology, we should shrug off many challenges to our basic logical and mathematical beliefs. But we cannot, as we should not, put aside challenges to our moral and aesthetic beliefs.

1 Abductive Arguments

Abductive arguments typically consist of a series of content premises, a “structural” premise which aggregates the information from the content premises, an abductive linking principle, and a defeasible conclusion drawn on the basis of the previous. The particular cases we’re interested in are comparisons of different theories where the conclusion tells us to accept one of these and not the other.

The point of articulating such an abductive argument is to *confer justification* on revising or rejecting our current beliefs in line with this conclusion. It’s this context that we focus on here, leaving questions about whether the resulting beliefs themselves are otherwise *justifiable* or *justified* to the side. This is a familiar notion of justification, albeit one that gets less attention than it should in the literature.⁵ This is especially striking since it’s this notion of justification that’s important for the case of revising our most basic beliefs about matters like mathematics, logic, morality, and aesthetics.

⁵See Engel (1992) for this point, sensible puzzlement about the lack of attention, and an analysis of this kind of justification.

I'll not worry overmuch about differences between wholesale revisions of our beliefs about these subject matters, rejecting our beliefs about them, and accepting wholesale doubt about them.⁶ After all, both rejecting and doubting all our beliefs about some subject matter are severe kinds of wholesale revision. We could explicitly represent them as such, but I'll simply trust the reader to not get confused.

Abductive arguments can be represented in many ways, including suppressing some of these premises into the act of inference itself. As it will make certain points easier to articulate, we'll be especially pedantic about representing premises and linking principles and "deductivize" them. We'll also make clear how to extend them by the claim that the defeasing conditions *are not met*, yielding the further conclusion that we ought to revise.

Our basic abstract argument compares the theoretical virtues of two incompatible theories of some phenomena (the generalization to n many theories is obvious):

CONTENT₁: Theory₁ of A has theoretical virtue P to degree n ,

CONTENT₂: Theory₂ of A has theoretical virtue P to degree k ,

...

STRUCTURAL Theory₁ is more theoretically virtuous than theory₂ on the basis of CONTENT₁, CONTENT₂, ...,

THEORY CHOICE: We ought to believe the most theoretically virtuous theory of A ,

SUB-CONCLUSION: We ought to believe theory₁ unless we have sufficient additional reason to believe theory₂;

UNDEFEATED: We do not have significant additional reason to believe theory₂,

CONCLUSION: We ought to believe theory₁ and not theory₂.

We can fill this schema to obtain particular abductive arguments.⁷ For example, let 'theory of A ' mean 'non-pragmatic explanation of our particular moral beliefs, intuitions, and perceptions'.⁸ Suppose that it's theoretically vicious for an explanation to posit systematic brute connections between facts. Then it's problematic to hold that our moral beliefs and intuitions are largely accurate while holding that there's no systematic explicable connection between our particular moral views

⁶These differences are important, but they're not important for the point I want to make below.

⁷This schema is just a working example, albeit an indicative one. The reader will easily be able to rejigger it for other uses of abduction.

⁸'Non-pragmatic' in the sense of a causal, grounding, or "real" explanation of our particular moral beliefs. I take the existence of non-pragmatic explanations for granted here.

and what makes them true.

After all, in the paradigmatic cases where we take our intuitions and perceptions to be accurate, we do have this connection. Our usual methods for explaining why we have the particular sensory perceptions we have *does* posit a causal connection between what makes the perceptions accurate and why we have them. So the connection is there non-brute. If the analogous connection is brute for some explanation of our moral beliefs, intuitions, and perceptions, then that's a strike against it.

So, absent some plausible explanation of this connection, seriously non-naturalistic pictures of morality are going to have a strike against them. Letting theory₁ of our particular moral beliefs be that they're the product of social conditioning and theory₂ of our particular moral beliefs that they somehow track real moral facts and properties where this tracking relation is itself brute, we get our content premises: theory₁ is relatively virtuous, theory₂ rather vicious.

If the rest of the virtues and vices of these explanations are roughly on a par, the weighing procedure is easy. By THEORY CHOICE, then, we've ought to adopt theory₁, rejecting theory₂, unless we have strong reason to hold onto theory₂. If we don't have sufficient additional reason to resist this conclusion (i.e. we have the relevant version of UNDEFEATED), then we ought to adopt theory₁ over theory₂.

Assuming we start off believing theory₂, we've just articulated a loose version of Harman's (1977) debunking argument against stark-raving moral realism.⁹ Whether Harman's argument is good is up for grabs; in particular, which theoretical virtues are relevant and what it would take to defeat SUB-CONCLUSION are very contentious, as I discuss in my (2018-a).¹⁰ Nevertheless, it's a clear instance of this kind of argument.¹¹

⁹I've gone with a simple "no brute connections" construal of Harman for simplicity. To reconstruct Harman's actual argument, we'd need to a lot more detail. For instance, he argues that the non-naturalist realist explanations targeted contain explanatory dross and then implicitly invokes the claim that such explanations are worse when there are explanations on offer without this dross (i.e. that compact explanations are better). Spelling this out takes a good deal of work. See (Sayre-McCord (1988). My (2018-a) contains a historically accurate reconstruction of Harman's argument along these lines.

¹⁰In fact, as I suggest there (see §3.1), much of the issue is what counts as a good reason to resist the challenge. Unsurprisingly, the typical targets of these challenges, non-naturalist realists, tend to be much more epistemically permissive than anti-realists, error theorists, and naturalist realists.

¹¹Slightly different debunking arguments are found in Street (2006) and Joyce (2001). These can also be put in something like the above form—which is a good exercise for the interested reader!

There are really two nested arguments here, indicated by the pair of horizontal lines. The intuitive conclusion of the “inner” argument—that we ought to believe theory₁ is defeasible; though here we’ve made the literal conclusion of in the inner argument indefeasible by listing the defeasing conditions in its antecedent. This is unusual, but it’s useful to forestall certain confusions.¹² This structure makes clear that it’s in light of UNDEFEATED that we can conclude outright that we ought to believe this or that theory. Since it will be useful for the below formulations, we’ll call the simple claim that we ought to believe this and not that theory the *directive* of both the “outer” argument to CONCLUSION and the “inner” argument to SUB-CONCLUSION.

We can rationally accept SUB-CONCLUSION without rejecting theory₂. To do so we just need justification for rejecting UNDEFEATED. This blocks moving to CONCLUSION and thereby blocks the obligation to revise. In fact, we’ll see below that it’s exactly the failure of certain inner abductive arguments against logic and mathematics to extend to outer arguments that insulates our logical and mathematical beliefs from certain challenges.

Our overall question is how seriously to take self-effacing abductive attempts to justify various epistemic states when the grounds for this justification would be undermined by adopting those epistemic states. We’ll thus focus on when we can knock back attempts to confer justification for wholesale revision or rejection of our beliefs about some subject matter, like the directive of our example demands.

1.1 Self-effacement

So what exactly is self-effacement? It’s when an argument undermines its own support in some way. To make this precise, we’ll define it with respect to the argument schema given above (trusting the reader to make the necessary changes for other cases.) An argument σ of that form is *self-effacing* when:

- We have sufficient support for all σ ’s premises preceding the first horizontal line;

See Schafer (2010), Vavova (2014), Lutz (2017), and Isserow (2018) for useful discussion of this Street-Joyce kind of evolutionary debunking.

¹²We’ve not exhibited the justification of each content premise. We could do this explicitly as well, but it would complicate things unnecessarily.

- Revising our beliefs about A in accords with σ 's directive¹³ would undermine our *actual* support for at least one of σ 's premises.

According to this definition, there can be self-effacing arguments whose conclusion we rationally accept. This is because accepting SUB-CONCLUSION means that our actual theory of A says of our actual theory of A that it's not the best and, therefore, that we've significant reason to reject it.¹⁴ This is true regardless of whether we can move forward to CONCLUSION. See also §1.5.

There are three ways for self-effacement to arise here:

- Revising our A -theory in line with σ 's directive would undermine our justification for CONTENT _{j} , for some j ;
- Revising our A -theory in line with σ 's directive would undermine our justification for STRUCTURAL or our understanding of one of the theoretical virtues used in CONTENT _{j} , for some j ;¹⁵
- Revising our A -theory in line with σ 's directive would undermine our justification for THEORY CHOICE;

We ignore the third way; it happens, but it's rarer and it's complicated.¹⁶ We'll call self-effacement arising the first way *content self-effacement* and self-effacement arising the second way *structural self-effacement*. We'll now give examples of each, working with the case of abductive reasoning about logical theories.

1.2 Logical Self-effacement

These examples are driven by the fact that some weak logics can't see their own virtues. For instance, the ability of non-transitive logics¹⁷ to prove claims of ordinary mathematics is informally "justified" using transitivity. But since transitivity

¹³That is, coming to believe theory₁ and rejecting theory₂, as suggested by SUB-CONCLUSION and demanded by CONCLUSION.

¹⁴I suppress here any argument that if we ought to reject some view, then we've got significant reason to reject it. Whatever view of reasons and ought you have, some such connection seems obvious. I'll leave it to the reader to situate my points in their favored framework.

¹⁵It wouldn't be unreasonable to divide this type of self-effacement into two, even though both have to do with the mechanics of abductive theory choice. I won't do so here for reasons of simplicity.

¹⁶So I won't discuss revising our commitment to the evaluative normativity expressed by the 'ought' in 'you ought to believe such and so'. But to put my cards on the table: such a notion of theoretical obligation is also a fundamental part of any cognitive project and is thereby immune from many challenges. There, I said it. See also Woods (2018-a, §5).

¹⁷See Tennant (2017) for the most rigorous version of such a logic and Cobreros et al (2012) for an interesting variation.

isn't a basic feature of a non-transitive logic, this is insufficient. Of course, since this recapture result is a bit of ordinary mathematics, if it's right there will be a non-transitive proof of it. Still it's not the kind of thing the non-transitive logician gets to take for granted—it needs rigorous proof, not a hand-wavy bootstrap.¹⁸

There are potential ways to defend the non-transitive bootstrapping approach, but they either require special pleading or are currently underdeveloped.¹⁹ We'll presume, along with the vast majority of folks working in technical fields, that we can only justify a bit of logical theory unless we can justify it, justify that we can justify it, etc. using logical resources we actually accept. Having noted this, here are examples of each type of self-effacement:

- **Content Self-effacement:** Suppose the target phenomena, our A , is logical consequence. Let theory₂ be our current transitive theory of logical consequence and theory₁ be a weak non-transitive logic. Since what a theory entails is beholden to the logic we use to suss out what it entails—to a background proof theory—changing our background logic can undermine our justification of the virtues of theory₂. It can undermine our content premises, in other words.

Take proving ordinary mathematical claims. This is a minimal virtue of a theory of logical consequence, but whether we can demonstrate that, say, the Peano axioms have their usual known consequences depends on having a strong enough proof theory in the background. The only sufficiently strong extant proof theory for non-transitive logics is done in classical logic.

We *can* demonstrate that theory₁ can recapture ordinary mathematical reasoning when we use theory₂ as our background logic. This underwrites a strong *prima facie* case for accepting theory₁. But rejecting theory₂ as a consequence undermines our justification for accepting CONCLUSION. Why? Because using the transitivity of consequence, as we can when we accept theory₂, is essential to current ways of showing that theory₁ recaptures ordinary mathematical reasoning. Changing from theory₂ to theory₁ on the basis of this fact would thus undermine our justification for the virtue of theory₁; that argument made use of transitivity. So we'd lose our justification for one

¹⁸This objection dates back to Burgess (2005). (Tennant 2017: 12.4) is his current rejoinder. See (Woods forthcoming-a) for why I think it's still insufficient.

¹⁹See (Woods forthcoming-a, forthcoming-b) for discussion. See also Meadows (2015) for useful criticism of similar bootstrapping approaches for the case of non-classical set theory.

of our content premises. So, such an argument, given the facts supposed, would be content self-effacing.

- **Structural Self-effacement:** Now let theory₁ be a logical theory so weak that it can't cleanly recapture elementary arithmetic. Forget about the content premises and focus on the corresponding version of STRUCTURAL. Justifying STRUCTURAL requires assigning weights to theoretical virtues like 'being able to recapture ordinary mathematical reasoning,' aggregating the various virtues and vices, then weighing it all out. If we hold that a theoretically virtuous explanation of *A* is a minimally plausible explanation of *A*, as we should, we also need to justify the minimal plausibility of theory₁.

All of this—comparing weights of virtues, explicating and justifying minimal plausibility, etc.—requires logic. If theory₁ is so weak that we can't do elementary arithmetical comparisons, we can't do any of it. So, if we were to adopt theory₁ on the directive of an argument like the above, we'd lose our justification for STRUCTURAL. Voila, structural self-effacement.²⁰

These cases establish that both kinds of self-effacement can happen. Of course, comparing the theoretical virtues of alternative logics and our own is a rather special case; we'll shortly look at a few slightly more humdrum cases.²¹ First, though, we need to briefly discuss the badness of self-effacement and put to rest a potential confusion.

1.3 What's Wrong with Self-effacement?

There's surely something problematic with justifying a conclusion by means of premises which would be undermined by accepting it. This is especially true for the purpose we've fixed: generating a justification *we can use* in justifying a revision of our current beliefs (see §1). But I shouldn't just thump the table and claim that:

BASIC FACT: It's irrational to revise our beliefs²² on the basis of conclusions where so revising would destroy the justification for that con-

²⁰For readers who doubt that some existing (and propagandized!) non-classical logics are this weak, I invite them to try to prove something non-trivial in elementary number theory, dotting every 'i' and crossing every 't', while using a non-classical logic weaker than intuitionistic logic, *without* invoking classical recapture. Or, alternatively, by proving classical recapture using this weak logic and *only then* invoking classical recapture.

²¹For excruciating details of the cases both described here and to be described below, see again Woods (2018-a, forthcoming-a, forthcoming-b).

²²I mean here for the revisions to be restricted to the right sorts of updating methods; I'm not worried about fancy "belief pill" sorts of cases for this reason.

clusion.²³

So here's a quick and dirty argument for BASIC FACT. Imagine how you'd justify your epistemic action to someone *after* acting on the basis of the conclusion of a self-effacing argument. You'd have to cite support which you no longer are entitled to. But you can't do that. Maybe you could cite support which is only now available from your new perspective. Even then you would have to view your prior *actual* justification, and thus your actual move to the new perspective from the old, as irrational. This situation seems epistemically vicious.²⁴

If self-effacement undermines rational retrospective endorsement of our epistemic action (holding fixed our actual reasons at the time of revision), it should also treat the putative prospective justification of revising as irrational. After all, in the cases we're imagining, we can easily see that even if we're entitled to the premises of our argument, we won't be after accepting and acting on the conclusion. So we're doomed, if we so act, to be unable to retrospectively endorse our epistemic action. And, given this, it seems reasonable to treat such epistemic actions as irrational.²⁵

So, on the basis of an intuitive "reflection" principle for belief revision, it's irrational to revise on the basis of self-effacing arguments since we can see in advance that doing so would undermine our justification for so revising. This *does not* mean that we're not currently entitled to the premises of a self-effacing argument; we sometimes are entitled to them, especially the ones preceding SUB-CONCLUSION. Rather, we're not justified to act on the conclusion of that argument, revising away our support. Since accepting CONCLUSION without revising or doubting our actual beliefs would be paradigmatically irrational, this means that we're not rationally

²³In our cases, we destroy the justification for the premises which entail the conclusion, but since justification is transitive in at least this case, BASIC FACT rules out rationally revising on the basis of self-effacing arguments.

²⁴There could be independent reasons to believe a theory which undermines our current justificatory standpoint; perhaps it's elegant, perhaps it's intrinsically intuitive, perhaps not doing so means putting up with a self-effacing theory and that's too high a cost to bear in a particular case (see §1.6 and §3). The point here is that we can't justify moving to that theory by using the undermined parts of our actual justificatory standpoint.

²⁵In my (forthcoming-a, forthcoming-b), I spotted the logical anti-exceptionalist that we can rationally revise from within a logical system even if our reasons for doing so essentially depend on our current logical theory. I was more generous in those essays since we can construct problems for revising logic abductively *even on this assumption*. Rewriting the current argument in a generous way would significantly complicate things. So I'll leave it to someone more generous than me to do so.

warranted in accepting the conclusion, even though the argument is valid.²⁶

After irrationally revising we may find *new* reasons to have acted as we have, but even then we can't justify why we acted as we did, though we can justify why we should have. So there's good reason to avoid accepting and acting on conclusions when so doing would undermine our reasons for accepting and acting on them. In particular, the presumptive truth of the materials necessary for our abductive investigations isn't open to doubt by means of these methods since this would undermine the project of using abductive methods at all.

1.4 The Self-Effacement Gambit

It's helpful here to draw an analogy with Wright's (2004) discussion of failures of warrant-transmission. Warrant-transmission was originally introduced to diagnose what's wrong with Moorean anti-skeptical arguments. Wright argues, contra Moore, that we can't gain justification or warrant for the belief that there's an external world from the output of the cognitive project of using perception to limn the world. This is because the existence of an external world is a presupposition of the cognitive project of using perception to limn the world. Given where that project starts, *that the world exists* is outside its remit; presuming that it does is required to make sense of perceptual warrant.

Inverting this, we cannot come to lose warrant or justification for a theory *T* *directly* (see §1.6) from the fact that skeptical conclusions about *T* can be justified *using T* or when engaged in *T*-involving cognitive projects. This is particularly acute in the case of *T*-involving abductive investigations into the best systematic account of our *T*-beliefs. Given what this project is and where it starts, systematic doubt about our *T* beliefs are outside its remit.

So this is our simple anti-skeptical strategy. When an argument is self-effacing, as many skeptical arguments turn out to be, then we're not rationally entitled to draw their conclusion since so doing would undermine our justification for their premises (by BASIC FACT.) Shortly we'll see that this simple anti-skeptical argument doesn't go far enough, but it does allow us to shrug off certain *particular*

²⁶Some of these arguments might be sound as well, though this is more contentious. The key question is how to do the costing in UNDEFEATED. In my (2018-a), I suggested that the self-effacement of debunking arguments against logic and mathematics was sufficient to provide a conclusive reason against revising our beliefs (blocking the move from SUB-CONCLUSION to CONCLUSION in that instance). This could be resisted by someone with a less conservative approach to justification. But either way, my earlier move conflated this move with a related one involving structural self-effacement.

arguments. That's not nothing.

Some will worry that if we can argue from within that we'd be better off with a different set of beliefs about some subject matter, then even if our argument depends on our current beliefs about that subject matter, we should nevertheless revise our current beliefs.²⁷ I think this worry comes from two mistakes.

One mistake is not distinguishing whether or not we can revise our beliefs in line with the conclusion of a *particular* self-effacing argument from whether we can revise our beliefs in line with the conclusion of a related non-self-effacing argument. We can often rejigger self-effacing arguments to avoid self-effacingness. For instance, arguments against particular logical principles often make use of that very principle. Usually, though, we can rerun the argument via other logical principles at the cost of some complexity. See below (§1.6) for another example of extracting a cogent argument against a particular subject matter from a self-effacing one.

The other mistake is a bit more impressionistic, but pervasive. There's a tendency to see arguments "dynamically", treating ourselves as moving from premises we accept to accepting a conclusion in a stepwise fashion.²⁸ This makes it seem as if we're moving stepwise from premises to SUB-CONCLUSION to CONCLUSION in a warrant-preserving process of belief updating. But actual belief revision doesn't work this way. When acting on CONCLUSION, we're actually accepting that accepting all the premises and the conclusion would be cogent (among other things!), *then* updating our epistemic state accordingly. A quick inspection shows that it's paradigmatically irrational to accept all the premises of a self-effacing argument, along with their support, while accepting its conclusion—it's tantamount to accepting '*p* and *p* isn't justified'.²⁹

²⁷Perhaps such theorists will also invoke the famed and asinine image of the ladder we kick away. But I've never understood how ladder-kicking could be justification-conferring, whatever other virtues it may have.

²⁸I suspect overfamiliarity with stepwise deductive calculi conjoined with underfamiliarity with actual mathematical proof is a large part of this mistake.

²⁹Strictly speaking there's a few moves in between accepting that we ought to revise our beliefs about a subject matter and the consequences of this for particular premises. Presuming this modest amount of closure seems unproblematic. I hope to explore elsewhere the general relationship of self-effacement to Moore's Paradox and other cases of pragmatic incoherence.

1.5 Problems with the Self-Effacement Gambit

We've been focusing here on the "outer" argument described above. Turn now to the "inner" argument. We can accept its premises without revising our beliefs about the subject matter or theory in question as there may be significant reason to avoid revising our beliefs about that subject matter. In other words, we can accept the SUB-CONCLUSION, deny UNDEFEATED, and thereby avoid drawing CONCLUSION. Is this a comfortable resting place?

No. Even accepting SUB-CONCLUSION without drawing CONCLUSION is problematic. Accepting SUB-CONCLUSION means holding that there's significant reason to revise our beliefs about some subject matter or phenomena. And, moreover, that these are demonstrable from within our current beliefs about that subject matter or phenomena.

Often we won't revise on the basis of these reasons because so doing would be problematic in other ways. But that's an uncomfortable position to be in; we're roughly accepting that our beliefs aren't entirely epistemically cogent from within. Even if we accept that revising would be severely problematic, we shouldn't be happy about our beliefs about some subject matter or phenomena are so down on themselves.

It helps to see the point if we distinguish between when a subject matter or theory is self-effacing and when a particular argument is. We've defined the latter above, but we can say that subject matter or theory is self-effacing when there are arguments it treats as good or valid which are self-effacing. Accepting an argument to SUB-CONCLUSION commits us to treating our standing theory of *A* as self-effacing and it's *this* which is the significant theoretical vice of *A*.

So, summing up, there's a strong case to be made that revising our beliefs in line with the conclusion of a self-effacing argument is irrational. It seems to violate an intuitive reflective principle for justification as well as being tantamount to accepting *p* and that *p* is unjustified simultaneously. This yields a strategy for blocking various challenges to our standing beliefs. Yet there are sound arguments whose directive (note 'directive', not 'conclusion') is revision of our standing beliefs about certain subject matters and phenomena. This very fact shows that there are problems with our accounts of various subject matters and phenomena; it underwrites reasons to revise, even if these reasons can be outweighed.

It's unclear how strong these reasons are; below I'll suggest that even though

we plausibly find such reasons in the context of logic and mathematics, they're not quite strong enough prop up challenges to our beliefs about logic and mathematics (at least fragment thereof, at least when we take abduction as our primary methodology.) I'll also suggest that when we turn to morality and aesthetics, the analogue reasons might be strong enough to challenge our standing beliefs. But, in order to see this difference, we need to do a bit more clarification.

1.6 What Self-effacement is not

Self-effacing arguments are easily confused forms of inference like *reductio ad absurdum*. On the basis of this confusion, some have objected there's no problem with self-effacing arguments since there's no problem with *reductio ad absurdum*. But this really is a confusion.

I suspect that the confusion comes a failure to cleanly distinguish *categorical* from *hypothetical* reasoning. Categorical reasoning involves moving from premises we accept to accepting conclusions which are supported by those premises.³⁰ Hypothetical reasoning involves reasoning *under a hypothesis* to establish facts about what follows from that hypothesis. Sometimes we mix the two up, such as when we infer that some proposition is mistaken on the basis of the fact that we can deduce an absurdity from it. This kind of inference involves *categorically* using a fact established by *hypothetical* reasoning—that φ yields absurdity—to conclude that φ is false.

There are ways of representing a rule like *reductio* where it looks like we come to reject a premise on the basis of a conclusion we categorically reach on the basis of that premise. But this is misleading. The most representative form of an actual argument by *reductio ad absurdum* isn't:

$$\begin{array}{c} \varphi \\ \psi \\ \hline \chi \wedge \neg\chi \end{array}$$

where we reject either φ or ψ on the basis of the fact that they jointly entail a contradiction. The most representative form is rather:

$$\begin{array}{c} [\varphi, \psi] \Rightarrow (\chi \wedge \neg\chi) \\ \hline \neg\varphi \vee \neg\psi \end{array}$$

³⁰Or moving from conclusions we reject to rejecting premises which support those premises.

where \Rightarrow represents deducibility under the hypotheses within the square brackets. Neither φ nor ψ functions as a premise in our reasoning by reductio; they are not part of the *categorical* considerations mustered in favor of our conclusion. Rather, the categorical consideration in play is the *deducibility of a contradiction from φ and ψ* .

Yet there's a grain of truth in most tempting objections, including this one. It's vicious for a theory to be self-effacing; to say of itself that it's not entirely correct.³¹ This is true *even if* we can resist revising our beliefs because the otherwise good argument that the theory is not entirely correct is self-effacing. So even though the analogy between reductio arguments and self-effacing arguments is incorrect, the fact that a theory generates self-effacing challenges to itself can sometimes play a role in a *different* categorical argument for revising our belief. This is rather analogous to how the deducibility premise plays a role in reductio proofs and so the temptation to run the two together isn't altogether surprising.

This new argument usually isn't self-effacing, it's rather an argument built on *top* of a self-effacing argument.³² As I'll shortly show, how problematic these sorts of arguments are depends on what would happen if we acted on them. We'll turn to this point next, but for now it's enough to recognize (a) revising our beliefs in line with the directive of a self-effacing argument is rationally problematic even though (b) the cogency of that self-effacing argument is itself a theoretical vice of our theory of *A*.

2 Locating Self-effacement

We've focused above on examples drawn from the theoretical side of rationality, in particular from logic. But nothing clearly prevents practical considerations from playing a role in whether a particular theory or explanation counts as best. Perhaps an explanation of my moral beliefs which has it that it's moral for me to have my moral beliefs is better than one which treats having those beliefs as immoral. Perhaps the goodness of a particular moral theory itself makes a difference. And so on.

If moral or even aesthetic considerations can play such a role, we open the door to challenges to morality and aesthetics being self-effacing. This is especially salient given the prevalence of debunking arguments against each. We'll take each case in turn, showing that self-effacement is possible in each case under certain

³¹See Schechter (forthcoming) for useful discussion of this point.

³²There may be cases where any such argument would itself be self-effacing. Set such cases aside.

hypotheses. We'll then show that we only get structural self-effacement for a small fragment of mathematics and logic; morality and aesthetics are, at most, content self-effacing.

2.1 Moral Self-effacement

While it's plausible that moral facts play a role in the best pragmatic explanation of our moral beliefs—I believe Roy Cohn was a monster because he was, in fact, a monster—it's controversial whether moral facts play a role in the best non-pragmatic explanation of our moral beliefs.³³ It's tempting to say that only physically realized phenomena like our *beliefs* about morality, the testimony of others about morality, and the like could explain our moral beliefs (Harman 1977).³⁴ After all, one standard defense of abductive methodology claims that the best explanation is the one most likely to be true. And, on many metaethical views, what makes facts about our psychologies true is independent from what makes facts about morality true.

On the other hand, aesthetic considerations like naturalness and elegance play a clear and plausibly non-instrumental role as theoretical virtues. This is perhaps most obvious for mathematical explanation, but it's easy to find instances in other contexts. So, unless we bizarrely treat prettiness as indicative of truthfulness, aesthetics plays a role in certifying certain explanations as the best. If aesthetic considerations can play such a role, why not moral considerations?³⁵

Some philosophers have argued, in particular, that the morality of a moral theory counts as evidence for its truth:

one morality is more likely to be true than another, because the former makes for a better world than the latter—not instrumentally, but intrinsically. (Nagel 1995: 92)

There being a difference between one's theory of the best normative X (the best morality, the best standards of inference, the best rules of justification...) and one's (so far) best theory of X, necessarily

³³See Sturgeon(1986) for the classic discussion of these cases.

³⁴Paulina Sliwa suggests (personal communication) that norms of truthfulness and trust like those undergirding testimony are an essential part of the explanation of our beliefs. She goes on to suggest that they thereby play a non-pragmatic role in explanations of our moral beliefs. Since the entanglement of these norms in our abductive methodology is quite deep, it would be a rather pressing case if the relevant norms were moral ones. I don't think the relevant kinds of truthfulness and trust are moral notions, but I don't have space to address the issue here.

³⁵See Hanson (forthcoming) for one set of reasons to treat these two subject matters analogously. See Woods (2018-b) for a distinctive set of reasons.

provides a reason (though perhaps not a decisive reason) to think one's (so far) best theory is wrong. (Sayre-McCord manuscript: 1)

Nagel's claim is stronger than Sayre-McCord's. Sayre-McCord's claim permits that it's indication of the truth of a theory that's doing the explanatory work. While on Nagel's view it's not merely a necessary consequence of our moral theory making for a better world that makes it more likely to be true, it's part of *why* it's likely to be true in a non-pragmatic sense.

To illustrate this difference, consider "third factor" views like Enoch's (2001) where moral beliefs and moral facts are explanatorily independent, both being instead explained by a common *explanans*. There's then a necessary connection between our beliefs about morality and the moral facts as they're explained by the same thing. This is consistent with Sayre-McCord's claim, but not Nagel's. It's not the fact that it's more moral that *makes* it the case that it's more likely true, but whatever the third factor is.

So there are strong and weak ways for morality to make a difference for which explanation or theory of morality is best.³⁶ It's even possible to avoid routing through the likeliness of a theory or explanation entirely, letting the morality of a view be an intrinsic part of its explanation. If we accept any of this, then the moral or immoral character of a moral theory can make a difference to what we ought to believe about morality.

Even the weak account is enough to generate self-effacement. Even if morality (or aesthetics) plays only instrumental role, our justification for the evaluative intermediate link can be undermined when coming to significantly revise our moral (or aesthetic) views. That's enough to produce cases of self-effacement, just as our inability to *see* or *prove* that a certain weak logic had certain undesirable properties was sufficient to generate logical self-effacement above. However, since I'm rather tempted to think that morality and aesthetics can play a more intrinsic role in making an account best, I'll not address instrumental roles in any more detail.

Now for self-effacement. Suppose that the second-best purely theoretical explanation of morality is that it's an effective fiction imposed by the ruling classes. That is, this Hobbesian view is the second-best explanation of morality once we *ignore any moral or practical virtues* or their effect on the plausibility of an explanation. Suppose, *per impossibile*, that some version of stark-raving (non-naturalist)

³⁶There's a litany of other defenders, such as Enoch (2009), Preston-Roedder (2014), and Quinn (1994). See Sayre-McCord for detailed and compelling defense of the argumentative strategy.

moral realism is a slightly better purely theoretical explanation.

Now suppose further that the Hobbesean view is a morally superior explanation of morality than realism. Perhaps it's morally better if we aren't beholden to mind-independent facts about how we ought to behave. Following Hume, maybe this is because it's morally best to have an explanation of morality that allows that we shouldn't be so slavish to moral concerns.³⁷ Or perhaps the type of moral realism in question requires us believe, on pain of immorality, that the immoral are intrinsically inferior; clearly that would be a pretty morally pernicious account of morality. Whatever the reason, if the Hobbesian explanation is sufficiently morally superior to believing in some kind of moral realism, it ought to take explanatory pride of place in our actual abductive reasoning about morality.

We can then argue that we should to reject moral realism on the grounds that we ought to believe the best (which includes the morally best!) explanation of morality. Yet the morally best explanation of morality is one which entails that our judgments about morality are systematically unjustified; they're actually capitalist fictions.

Since coming to reject our actual moral beliefs would immediately undermine our justification for this rejection—since part of why the Hobbesian view was the best overall explanation was its *moral* character—we have a case of content-based self-effacement. The example is merely illustrative; we could argue like this in several different ways.³⁸ It's enough here to demonstrate plausible ways such an argument could go.

Morality can only be self-effacing in the content way though—once we've settled the virtues of a good explanation or theory, there's no further *moral* question which arises in explicating these virtues or aggregating them. We'll draw on this below to argue that there's still a determinate sense in which morality is more vulnerable to the kind of challenges we're concerned with, even though it permits self-effacement. But first aesthetics.

³⁷See Schafer (2016) for a useful explication of Hume on this point.

³⁸We could also, for instance, use claims about the essential role of agency in our cognitive projects in place of the Nagel or Sayre-McCord strategy. Or we could make use of Hayward's (forthcoming) recent defense of the immorality of moral realism.

2.2 Aesthetic Self-effacement

Whether or not morality partially determines the quality of some explanations, aesthetics surely plays such a role. *Elegance*, *naturalness*, and *non ad-hoc-ness* are often treated as important virtues of explanations and theories.³⁹ This is most glaring in mathematics as noted by Hardy (1967), but it's also clear on inspection that we use these virtues in nearly all actual applications of abductive machinery.

Similarly to moral considerations, aesthetic considerations might be relevant because they evidence the explanation most likely to be true. Perhaps the world is more beautiful than it appears. Perhaps it won't tolerate ugly explanations or theories—the beautiful can be really horrid that way. We need not treat aesthetic considerations as so flatly instrumental though.

Lipton (2004) treats 'loveliness', explicated in terms of ability to convey understanding, as a primary theoretical virtue, albeit one connected non-trivially to likelihood. This is clearly a view on which aesthetic virtues play an important instrumental role in justifying that some explanation is the best, albeit a more complicated one than the version just mooted. He goes on to cite the ability to capture the aesthetic virtues of theories in terms of loveliness as part of the attraction of 'inference to the loveliest explanation'.

We might also flat out accept that the best explanation of some phenomenon isn't independent from the prettiest, most natural, and least ad hoc explanation. So, as with the discussion of morality above, we can formulate stronger and weaker roles for aesthetics to play. Each role seems to be *prima facie* reasonable and familiar from our actual abductive practice.

Yet the virtuousness of aesthetically pleasing explanations is more independent of the subject matter than the virtuousness of morally pleasing explanations. While a morally better explanation of morality is perhaps a better explanation of morality, a morally better explanation of some physical phenomenon is by no means a better explanation. Besides lacking the pithy expression, it's simply seems irrelevant to the explanatory goodness of physical explanation or the virtues of a physical theory whether or not it conforms to our moral scruples.⁴⁰

³⁹See Lipton (2004: 66) where these virtues are explicitly labeled aesthetic.

⁴⁰What we should *believe* about physics is perhaps a different story, but even here, it's *prima facie* implausible that morality will play a potential *significant* role like it might for our moral beliefs. See Maguire and Woods (manuscript) for a survey of cases.

Of course, not all of aesthetics plays a role in explanation. It simply irrelevant to the explanation of a particular mathematical theorem whether the explanation is funny. But it seems hard to deny that the theoretically-oriented aesthetic virtues are relevant; after all, it's a constant source of interest in mathematics and physics to show that a particular explanation is not only plausible, but also elegant, simple, and natural.

So at least the theoretically-oriented aesthetic virtues listed above plausibly play a role in abduction regardless of the subject area of the explanandum. We have a contrast here between whether the virtues of a particular subject matter are sensitive to the subject matter of what we're aiming to explain or not. Morality yes. Aesthetics, or the theoretically-oriented fragment thereof, no.

Aesthetic properties also only give rise to cases of content self-effacement (we'll not construct such cases here; just follow the recipe above.) After all the virtues are in and weighed, there's no further important aesthetic question about which theory wins out.⁴¹ It's also unlikely that aesthetic facts play a significant role in explicating particular virtues, aesthetic or no, in the way that mathematics and logic explicate notions of prediction, commitment, and probability. So aesthetic self-effacement seems restricted to content self-effacement.

2.3 Logical and Mathematical Self-effacement

We've given a few particular cases of self-effacing arguments against our current logical theories above, so we'll not develop more particular examples. The reader won't have trouble generating them on their own either since mathematics and logic play an especially entangled role in our abductive reasoning: both justifying and making any sense of typical theoretical virtues involves making use of something like a theory of entailment and a minimal fragment of arithmetic.

Consider theoretical strength. This virtue⁴² can—in fact, must—be explicated in terms of what a particular theory predicts. And prediction is at least largely explicated in terms of what a theory entails. Theoretical simplicity needs a notion of complexity which is typically developed in a fragment of arithmetic. Ontological simplicity needs a notion of cardinal comparison which itself requires a significant

⁴¹There's an aesthetic question for any actual calculation. We shouldn't bobble around when we can calculate cleanly and efficiently. But that's not relevant to which explanation is more virtuous than which.

⁴²Though Russell (forthcoming) suggests that theoretical strength isn't clearly a theoretical virtue.

chunk of either mathematics or logic.⁴³

Even weighing out various virtues is done in a background where we assign relative weights to these virtues, then aggregate them. This involves a minimal theory of aggregation, which isn't entirely trivial.⁴⁴ Finally, in order for an explanation or theory to be best in the relevant sense for abductive justification, it shouldn't just be the best of the worst. It should also be minimally plausible. This seems to require the notion of being minimally probable. Which, of course, is theorized mathematically.

So, unlike in aesthetics and morality, we can have both content and structural self-effacement for challenges to our logical and mathematical beliefs. This is what makes the case of abductive challenges to logic and mathematics distinctive (though see fn. 29.) We'll focus on structural self-effacement since the content case resembles the above discussions of logic and mathematics.

We don't need all of mathematics and logic to employ abductive methodology. We don't even need much mathematics and logic. What we need is a minimal fragment of theoretical power from some combination of mathematics and logic with which to compare various explanations.⁴⁵ Let me explain.

We have a choice when developing our account of these background notions. We can work in a small fragment of mathematics, such as a fragment of arithmetic, or we can work in a suitably strong background logical theory. For example, we could situate the theory of cardinal comparisons in a logic augmented with a binary quantifier $M(X, Y)$ expressing that there are more X s than Y s. Alternatively, we could define the relevant notion, on minimal background assumptions, in second-order logic. We could also just define the relevant notion in first-order set theory. Any of these methods and many others beside would suffice, though obviously each has virtues and vices. What we really need is *some* system with the appropriate amount of theoretical strength. It doesn't matter so much what it is.

⁴³Finite cardinal comparisons, of course, can be done in an insignificant chunk of logic. Interesting theories are rarely finite.

⁴⁴I'm skipping the justification of aggregation functions since we're now flogging a dead horse. Adding this in would add yet another place where appeal to a minimal fragment of mathematics and logic is necessary.

⁴⁵For the cognoscenti, what we need is a theory which contains something like a theory of inductive definitions, the ability to do arbitrary cardinal comparisons, etc. It would be shocking if our needs exceeded the logical strength of a very weak subsystem of arithmetic, such as WKL_0 .

Correspondingly, we only get structural self-effacement of an interesting sort when we formulate a challenge to the truth of *some minimal fragment or other*. We cannot get a direct justification for holding onto any particular way of developing this fragment since we can always use a distinct way of developing this fragment to run an abductive argument against the first way (exercise for the reader).

So we can't exploit self-effacement to justify any *particular* mathematical and logical beliefs against reasonable challenges. What we can do is justify the existence of a small fragment, of whatever type, which would serve to do the work we need for abductive reasoning. This means that we're not defending our actual logical and mathematical beliefs so much as rejecting wholesale skepticism about the resources with which to carry out abductive investigations. And this is a good thing! It would be incredibly dogmatic to defend particular sets of logical or mathematical beliefs against reasonable objections on these kinds of grounds.⁴⁶

It's only because these fragments share a common amount of "theoretical power" that we can use any one of them to challenge another. But morality and aesthetics don't plausibly share this property; there's no common core of moral or aesthetic theory that's strictly necessary for abduction. And so there's a striking way in which these latter two cases of self-effacement differ from the logical and mathematics case.

Summing up, accepting that there's some minimal fragment or other that suffices to carry out the *machinery* of abductive reasoning is necessary for any actual abductive reasoning—after all, without it, we couldn't make sense of or justify premises like STRUCTURAL. In contrast, for both morality and aesthetics, we could revise our abductive practice without abandoning abductive methodology entirely.⁴⁷

3 Types of Self-Effacement

We've now seen a number of cases of self-effacement covering a number of different areas. We can summarize the epistemically relevant differences between these as follows:

⁴⁶There are cases of content self-effacement that allow the defense of our particular logical and mathematical beliefs. These are not difficult to construct using the recipes already given above.

⁴⁷It's strictly possible to pull a similar generalizing move for morality and aesthetics, claiming we only need something capable of fulfilling some functional role to underwrite certain challenges. Yet it's not at all clear that this functional role must be fulfilled for abductive reasoning, so we still don't have structural self-effacement.

- Morality permits content self-effacement, but morality’s role in abduction is context-sensitive. Most applications of abductive reasoning don’t make use of moral content, even if abductive reasoning about morality does. We also need not salvage morality for a recognizable notion of abductive methodology, even one applicable to morality. So even if morality plays a role in determining which explanation or theory of morality is best, this isn’t essential.
- Aesthetics also permits content self-effacement, but aesthetics’ role fairly in abduction is context-independent. Still, we need not salvage aesthetics for a recognizable account of abductive reasoning. We could get rid of our affection for pretty theories without enormous loss, even though the resulting abductive methodology would look pretty different from our actual abductive methodology.
- Logic and mathematics permit both content and structural self-effacement and they play a context independent role in abduction. This is because we require some common “amount” of mathematics and logic to explicate and compare theoretical virtues for any application of abductive reasoning.⁴⁸ That is, logic and mathematics are required for any recognizable notion of abductive methodology as applied to any context.

3.1 Weighing the (theoretical) Costs

These differences mean that the theoretical costs of acting on the directive of a self-effacing argument differ quite heavily between different types of self-effacement.⁴⁹ Remember here that these costs are reckoned from our *current* standpoint; we’re costing wholesale, not piecemeal, revision or doubt of our beliefs about mathematics, logic, aesthetics, or morality given what we currently believe about these domains. So even if revision would leave some anemic remnant in its wake, that’s close enough to abandonment for our purposes here.

⁴⁸Some particular arguments might involve materials which are context-dependent as well, but that’s not important for the points below.

⁴⁹Again, I’m focusing on the *theoretical* costs of wholesale revision of our moral, aesthetic, logical, and mathematical beliefs. There are also *practical* costs which will weigh in on some views. Since it’s a contentious issue how to weigh practical and theoretical costs and since we only need theoretical costs for the defense of logic and mathematics I’m going to offer shortly, I’ll put these issues to the side. It’s enough to see that there is a question for morality and aesthetics about how costly wholesale doubt would be—though, as many error theorists have shown, it’s not nearly as high as people seem to think.

In the case of moral self-effacement, the costs still don't seem especially high. Even though we might have to revise our actual abductive practice once we revise morality, our successor notion wouldn't be wildly far away. After all, even if we come to systematically doubt, reject, or massively revise our moral beliefs, presumably we'd stop treating morality as a theoretical virtue of moral explanations and theories. And this is a relatively minor revision of our abductive machinery, or so it seems to me. So, even considering the worst theoretical costs, self-effacement isn't enough to protect morality.

Purging our abductive methodology of aesthetic virtues would amount to a quite dramatic overhaul of our abductive machinery. We use theoretically-oriented aesthetics constantly in evaluating explanations and theories. Junking them or significantly revising our aesthetic views would thus amount to an enormous change to our abductive practices. And this seems rather costly. So weighing out the reasons for and against a challenge to our aesthetic beliefs, it's not clear what to do. But potentially we might decide that this revision of our abductive methods isn't costly enough to outweigh the fact that our best abductive methods instruct us to wholesale revise, reject, or doubt our aesthetic beliefs.

The effects for structural logical and mathematical self-effacement would be catastrophic;⁵⁰ we cannot escape using some fragment of logic and mathematics when engaging in abductive reasoning. So, at least for the minimal fragment thereof that we've been interested in, it *is* clear that we've sufficient reason to resist revising.

To sum up, abductive challenges to all our logical and mathematical beliefs threaten not just a revision of our abductive practices when applied to logic and mathematics, but to our practices generally and thus to the cognitive project of applying abductive methods in coming to form our beliefs about the world. This isn't true of aesthetics, much less morality. Consequently, we have sufficient reason to resist revising for logic and mathematics, less for aesthetics, and even less for morality.

3.2 Persistent Worries about Morality and Aesthetics

Returning to our original argument schema, this means that the claim that we don't have sufficient reason to resist revising—UNDEFEATED—is most plausible for the

⁵⁰The case of content self-effacement for logic and mathematics is trickier. We put it to the side for now, noting that Woods (2018-a, forthcoming-a, forthcoming-b) address closely related issues.

moral instance of the argument, less so for the aesthetic instance, and clearly implausible for the logical and mathematical instance. Correspondingly, we cannot extend the logical and mathematical instance of SUB-CONCLUSION to CONCLUSION, though we potentially can for morality and aesthetics.

By BASIC FACT, we cannot get warrant for revising our beliefs about morality and potentially aesthetics from these self-effacing arguments themselves. But, as I noted above, if a subject matter *A* endorses self-effacing arguments, it says of itself that it's not the best. In fact, as we can potentially move from SUB-CONCLUSION to CONCLUSION in the moral and aesthetic instances of our argument schema, it's not just that we have some reason for wholesale revision or rejection of our standing moral and aesthetic beliefs. We, in fact, *ought* to do so. Yet doing so wouldn't be justified on the grounds justifying that we ought.

This further consequence is so bad, perhaps, that the fact that our current moral and aesthetic beliefs yield such arguments gives us *independent* reason to reject our moral and aesthetic beliefs. This is an even stronger form of the challenge gestured at in §1.6. Even without this, we may have sufficient reason to revise simply on the grounds that our moral and aesthetic beliefs are self-effacing—that is, on the grounds that they permit a good argument to SUB-CONCLUSION.

As pointed out above (§1.6), the argument just sketched shouldn't be conflated with self-effacing argument itself. It's rather a different argument against our standing moral and aesthetic beliefs which uses the fact that our moral and aesthetic beliefs permit self-effacement as a premise. We need to weigh the costs of revising our moral and aesthetic beliefs against the costs of sticking to our guns and tolerating good self-effacing arguments. It's not immediately clear what the outcome of this weighing will be, though I suspect it looks pretty bad for morality and aesthetics. But the potential for challenges of this type exists, whatever the ultimate outcome of the weighing is.

This new challenge cannot be run against our logical and mathematical beliefs. We have *pro tanto* reason to revise these beliefs, sure. But our reason to revise is easily outweighed by the catastrophe of losing our ability to use abductive arguments at all. As we cannot move from SUB-CONCLUSION to CONCLUSION in this case—given the catastrophe of revision—we don't get the more vicious conclusion either. So we seem to have sufficient reason to resist revising. And this, in turn, is the hull of the revised anti-skeptical argument I mentioned at the outset.

4 Conclusion: The Revised Anti-Skeptical Gambit

Drawing the above discussion together, our revised “abductivized” anti-skeptical gambit goes as follows. For the privileged fragment of logical and mathematical beliefs we’re concerned with, we cannot reason to CONCLUSION since there’s overwhelming reason to not act on SUB-CONCLUSION. This means that we’re ultimately just weighing *pro tanto* reasons to revise our faith in some minimal fragment of mathematics and logic against the catastrophe of abandoning abduction altogether. This is a calculation whose outcome is easy to see in advance.

The corresponding calculation for morality and mathematics is much less clear since revision doesn’t bring (theoretical) catastrophe in its wake. So we have no robust defense against the challenges to morality and aesthetics arising from the existence of good self-effacing arguments, even though we can’t use those self-effacing arguments to reject morality and aesthetics directly.

Returning to our earlier mention of Wright (2004), we can put the point another way. It’s a presupposition of being a cognitive project at all (in the relevant sense) that enough of our background mathematical and logical beliefs are correct. Rejecting them is rejecting the possibility of such cognitive projects. This is far too high a cost to be counterbalanced by the irritating fact that our logical and mathematical views support arguments that they’re not the best views to have. And we can thus put off challenges built on this irritation. But morality and aesthetics are not part of what it is to be a cognitive project at all, so they’re not similarly protected.

It would be fascinating if, in contrast to what I’ve argued, aesthetics or morality could be shown to play a deeply entangled role in our abductive methodology analogous to that played by mathematics and logic.⁵¹ If I’m right that their involvement is shallow, then there’s a deep divide in how vulnerable areas like logic and mathematics and less foundational subject matters like morality and aesthetics are to abductive challenges. If I’m wrong, then there’s a novel anti-skeptical defense available to us that holds for aesthetics and morality as well as logic and mathematics—just riff off the argument I gave above defending logic and mathematics. Either way, paying attention to the role our theories of morality, aesthetics, logic and mathematics play in abductive reasoning reveals quite a bit about what we are and are not in a position to take a skeptical swipe at. And this, I reckon, is a good thing.

⁵¹Though see fn. 22 above.

References

- Burgess, J. P. 2005. No Requirement of Relevance. in *The Oxford Handbook of Philosophy of Mathematics and Logic*, ed. S. Shapiro, Oxford University Press: 727–750.
- Cohen, G. 2000. *If You're An Egalitarian, How Come You're So Rich?* Harvard University Press.
- Enoch, D. 2009. Wouldn't it be nice if p , therefore, p (for a moral p). *Utilitas* 21(2): 222–224.
- Enoch, D. 2011. *Taking Morality Seriously: A Defense of Robust Realism*. Oxford University Press.
- Hanson, L. 2018. Moral Realism, Aesthetic Realism, and the Asymmetry Claim. *Ethics* 129(1): 39–69.
- Hardy, G. H. 1967. *A Mathematician's Apology*. Cambridge University Press.
- Harman, G. 1977. *The Nature of Morality: An Introduction to Ethics*. Oxford University Press.
- Hayward, M. K. forthcoming. Immoral Realism. *Philosophical Studies*: 1–18.
- Hjortland, O. T. 2017. Anti-Exceptionalism about Logic. *Philosophical Studies* 174(3): 631–658.
- Isserow, J. forthcoming. Evolutionary Hypotheses and Moral Skepticism. *Erkenntnis*: 1–21.
- Joyce, R. 2001. *The Myth of Morality*. Cambridge University Press.
- Lewis, D. K. 1983. *Philosophical Papers*. Oxford University Press.
- Lipton, P. 2004. *Inference to the Best Explanation*. Routledge.
- Lutz, M. forthcoming. What Makes Evolution a Defeater? *Erkenntnis*: 1–22.
- Maguire, B. and J. Woods. manuscript. Explaining Epistemic Normativity.
- McPherson, T. 2015. The Methodological Irrelevance of Reflective Equilibrium. In *The Palgrave Handbook of Philosophical Methods*, ed. Christopher Daly, Palgrave Macmillan: 652–674.
- Meadows, T. 2015. Unpicking Priest's Bootstraps. *Thought: A Journal of Philosophy* 4(3): 181–188.
- Nagel, T. 1995. Personal Rights and Public Space. *Philosophy and Public Affairs* 24(2): 83–107.
- Preston-Roedder, R. 2014. A Better World. *Philosophical Studies* 168(3): 629–644.

- Priest, G. 2006. *Doubt Truth to be a Liar*. Clarendon Press.
- Priest, G. 2016. Logical Disputes and the *a priori*. *Logique et Analyse* 59(236): 347–366.
- Quinn, W. 1994. *Morality and Action*. Cambridge University Press.
- Russell, G. 2015. The Justification of the Basic Laws of Logic. *Journal of Philosophical Logic* 44(6): 793–803.
- Russell, G. forthcoming. Deviance and Vice: Strength as a Theoretical Virtue in the Epistemology of Logic. *Philosophy and Phenomenological Research*.
- Sayre-McCord, G. manuscript. On a Theory of a Better Moral Theory and a Better Theory of Morality.
- Sayre-McCord, G. 1988. Moral Theory and Explanatory Impotence. *Midwest Studies in Philosophy* 12(1): 433–457.
- Scanlon, T. M. 2014. *Being Realistic About Reasons*. Oxford University Press.
- Schafer, K. 2010. Evolution and Normative Scepticism. *Australasian Journal of Philosophy* 88(3): 471–488.
- Schafer, K. 2016. The Modesty of the Moral Point of View. In *Weighing Reasons*, eds. E. Lord and B. Maguire, Oxford University Press: 241–256.
- Schechter, J. forthcoming. Is There a Reliability Challenge for Logic? *Philosophical Issues*.
- Street, S. 2006. A Darwinian Dilemma for Realist Theories of Value. *Philosophical Studies* 127(1): 109–166.
- Sturgeon, N. L. 1986. Harman on Moral Explanations of Natural facts. *The Southern Journal of Philosophy* 24(1): 69–78.
- Tennant, N. 2017. *Core Logic*. Oxford University Press.
- Vavova, K. 2014. Debunking Evolutionary Debunking. *Oxford Studies in Metaethics* 9: 76–101.
- Williamson, T. 2014. Logic, Metalogic and Neutrality. *Erkenntnis* 79(2): 211–231.
- Williamson, T. 2017. Semantic Paradoxes and Abductive Methodology. In *Reflections on the Liar*, ed. B. Armour-Garb, Oxford University Press: 325–46.
- Williamson, T. manuscript. Abductive Methodology and Applied Mathematics.
- Woods, J. 2018-a. Mathematics, Morality, and Self-Effacement. *Noûs* 52(1): 47–68.

- Woods, J. 2018-b. Footing the Cost (of Normative Subjectivism). In *Methodology and Moral Philosophy*, eds. J. Suikkanen and A. Kauppinen, Routledge: 166–190.
- Woods, J. forthcoming-a. Logical Partisanship. *Philosophical Studies*.
- Woods, J. forthcoming-b. Against Reflective Equilibrium in Logical Theorizing. *Australasian Journal of Logic*.
- Wright, C. 2004. Warrant for nothing (and foundations for free)? *Aristotelian Society Supplementary Volume* 78(1): 167–212.