Deep Disagreement, the Owl of Minerva Problem, and Pathologies of Meta-Argumentation

Scott Aikin

Abstract
This essay is an overview of the Owl of Minerva Problem – that our conceptions of argumentation and the practice of reason-giving, when looped into our practices and reason-giving as new arguments about our arguments (meta-arguments), can not only improve our performances but also warp them. Call these negative loops of reflection on argument *pathologies of meta-argumentation*. Exemplary of this phenomenon is the fallacy fallacy. Proposed solutions to the Owl of Minerva Problem from David Godden (2022) and Andrew Aberdein (2022) are addressed, and the problem is then shown to have application in how we conceptualize deep disagreement. Pathological meta-arguments about deep disagreements are termed *deep disagreement fallacies*. These meta-argumentative pathologies obtain only for particular conceptions of deep disagreement, but they are not seen to problematize the *dialectical depth* conception of deep disagreement, which accounts for depth in terms of the variety of inferences and the embeddedness of the commitments constituting the controversy. This is a pragmatic reason to prefer the dialectical depth conception of deep disagreement.

1.

Consider the following reasonable hope. We have practices with norms, and those practices can be improved with activities of making those norms explicit to practitioners. This activity of explicitification will, we expect, occasion better norm-compliance, enforcement, and perhaps even encourage some norm-improvement, too. That is, if we make the norms of our practices explicit, we allow ourselves the opportunity to intentionally follow these directives, and we can, for the sake of correction, call attention to how others fall short. Further, we can even tinker with the norms, perhaps asking how strictly to enforce them, whether there should be exceptions, and whether these norms

are irreducibly plural or are part of one big objective of the practice. Thereby, we hone our shared lives together, and we, in this reflection, come to understand ourselves better. This big hope is the clarion call of much theory and simple reflection, whether it is in ethics, sports, institutional design, research practices, or just in a loving relationship. The originating thought is that in explicitifying the norms and ends constituting a practice, we can engage in the practice better, perhaps improve it, and see it all in the larger picture of our lives together. And this is the teleology of informal logic, critical thinking, and argumentation theory – in this case, the practice is that of arguing about things with others (and perhaps ourselves), and the hopes for improvement are those of making our given arguments better, ensuring that our exchanges of arguments are more productive, recognizing tempting but common errors, finding paths to stable resolution, and understanding (and endorsing) the practice as part of our common intellectual life. Hey, that's not nothing. In fact, that's quite a lot!

The general objective is that we want to have this positive loop between our reflection and our practices. We do and say things, and on reflection, we think about those doings and sayings. We name particular errors, moves, successes, and we theorize them to explain and evaluate them. Further, we see room for improvement of our norms. Then we bring those results of reflection back to practice, perhaps with the intervention of some coaching, a lesson in theory, an institution constraining the practice, or enforcers of codified rules. With sports, we've got referees, and they are supposed to not only be intimately familiar with the details of the rules, but they are supposed to enforce them without favor. The problem for many other practices is that rule enforcement is a matter often for the participants, and if the practice has some division in it, it falls to partisans. We can see why that's a challenge, on an analogy returning to sports. *Offsides* in soccer (yes, I'm an American) are contentious enough calls when you don't have standing worries about the motives of the person holding the linesman's flag. But it's virtually impossible to feel right about it when it's a center back who's also running the trap with his arm in the air. Or consider the *dive* in the penalty box – without a non-partisan person making that judgment, penalty calls will be nearly impossible to enforce without wild escalation. Partisanship and rule-enforcement have a problem, because invoking the rule for one's benefit can be part of a game plan.

My target area of concern, again, is that of argumentation. We argue about things all the time. What to make for dinner tonight, whether the dog needs a bath, which Finnish black metal band is the scariest, and why democracy is facing problems right now. The key is that as we argue about these practical, objectual, and conceptual matters, we also argue about the arguments. Why is

our neighbor's veganism relevant to our dinner choice? Is the analogy between our dog and a human child apt when it comes to hygiene? Why can't *most Satanic* be the prime criterion for band scariness? And what's wrong with comparing real and failing democracies with ideal totalitarian alternatives? Plato did it, and if it's good enough for him…. The point is that we see patterns, errors, innovative connections and processes for thinking it all through, and we not only keep one eye on what we are arguing about in the first place, but we also are attentive to how we are doing it. As we are arguing, we are also open to further meta-arguments, arguments about the arguments' strength, relevance, appropriateness, and so on. So, when you say *therefore* or *consequently*, you're not just announcing that you have arrived at a conclusion, you're calling attention to the fact that you've come to it in a fashion that should not only invite but positively constrain my acceptance, because it is part of a shared practice designed for the purpose of dispute resolution and belief refinement.[1]

Argumentation theory and informal logic comprise a discipline of studying not only the norms of argument-exchange but developing interventions to improve practice. So, there is a longstanding connection between the informal logic movement and critical thinking initiatives. I teach a critical thinking class at least once every academic year, and the course satisfies the university's liberal arts requirement. It's not a surprise, really, as the requirement for the class is the interventions-edge of the hope of positive looping between theory and practice. Students show up with some basic reasoning skills, we reflect on them, explicitify a few principles, endorse some rules, and practice identifying good moves and calling out fallacies. And then students have their reasoning improve – that's the aspiration of the requirement and the class. But problems emerge quickly. The most troubling, given my purposes here, is how often I see students become fallacy-hounds, instead of better and fairer reasoners.[2] The tools of critical thinking become weapons in their hands, or at least new tools for argumentative confusion. The students are more like sharks or hyper-critical naysayers with a specialized vocabulary to use, interpreting every mention of another's inconsistency as *ad hominem tu quoque*, invocations of an authority as *ad verecundiam*, and any concern about the moral costs of an error as *ad misericordiam.* All that Latin was supposed to be a meta-linguistic tool to pull us out of the arguments on the first order to survey their quality on a second-order. But it became a powerful tool for obfuscation and obstruction. That's ironic, right? Well, I don't just think it's ironic, but it's a real problem for that reasonable hope we'd opened with.

---

[1] See (Pinto, 2001; Casey, 2020; McKeon, 2021, 2022; Aikin and Casey, 2024).
[2] See Edward Damer's similar observations of 'fallacy mongering' (2012).

One particular phenomenon worth a deeper look as an object-lesson in this troubling arc of reflection is the *fallacy fallacy*. It comes in a few forms, but it all starts with the fact that we have a participant in a critical dialogue using a meta-linguistic and meta-argumentative instrument in the midst of the dialogue. In the *inferential form*, the fallacy fallacy is inferring that since an interlocutor's argument is a fallacy, its conclusion must be false (Cotton, 2018; Aikin and Talisse, 2020). In the *attributional form*, the fallacy fallacy is the over-identification of fallacy to arguments that are in fact appropriate (Miller and Miller, 2015; Hedden, 2019). And in the *disagreement form*, the fallacy fallacy is the starting assumption that those who disagree must have arrived at their conclusions by fallacious inference, so all exchange is performed with the objective of identifying the fallacy to attribute the disagreement to (Lycan, 1996). The point is that with the vocabulary of fallacies, we've identified common errors, but instead of using those concepts to improve the discourse, we've developed new pathologies in the discourse, ones that would be otherwise impossible. Let us call these *meta-argumentative pathologies,* problems with argumentation that emerge because we are using the meta-argumentative concepts of correction in ways that are at cross-purposes with the practice.

Notice that this meta-argumentative error-type of the fallacy fallacy is one that extends beyond the simple form of fallacy-identification or accusation.[3] The procedural rules of argument, those of freedom of expression, balances of reason, requirements of reply, and norms of negotiation of what the proper controversy is, can also be turned into accusations of errors. The *free speech fallacy* is a meta-argumentative pathology that requires that challenges to views be given consideration and reply without restriction; *whataboutism* is the challenge to whether an item for scrutiny is appropriately evaluated given other similar but not discussed matters, and *bothsiderism* is the meta-argumentative inference from the fact of controversy to the conclusion that the issue is more complex than anticipated and so requires all, even the apparently irrational, voices in the conversation (Aikin and Casey, 2023, forthcoming). Add to these phenomena the thought that our vocabulary of critical thinking can easily serve as a menu for straw manning one's opposition, we see the problem clearly (Aikin and Casey, 2022).

The reasonable hope of positive looping between practice and theory has success cases for sure, but it also has other troubling consequences: these meta-

---

[3] For further accounts of meta-argumentative norms, see (Cohen, 2001; Finocchiaro, 2013; Wohlrapp, 2014). Further, see (Linker, 2014; Innocenti, 2022; Stevens, 2022) for accounts of abuses at the meta-level.

argumentative pathologies. And they are troubling because they are instances of where the norms and their explicitification have been instrumental in making the practice worse. There is a tragic tinge to our initial hope.

2.

The Owl of Minerva flies only at dusk. Hegel famously notes this in his *Philosophy of Right*, as he complains of how reason comes too late to the problems it should predict and prevent, were it prospective. Reason allows us to explain the errors after we make them, and we can resolve to do better. But philosophy, painting grey on grey with its fine distinctions and technical vocabulary, recedes from the scene as a mover. On this Hegelian line, philosophy arrives too late, but even if right, makes little difference. The dark Hegelian view is probably right, but I think there's an extra problem – the explanations and normative resolutions are sites for further and new errors. And, again, with argument, we see this with the fallacy fallacy and pathologies of meta-argumentation. Reason shows up late, with explanatory and normative tools to set things right. But because those tools are put in the hands of partisans and in the hands of those who needed correction for their reasoning in the first place, their corrections to work which was originally badly done will itself also be badly done. We are fallible creatures, and implementations of measures to correct for our fallibility, we should expect, will be fallible and fallibly carried out. So even if we have the tools of wisdom to make us more wise, we, because we are not yet wise, will misuse, misunderstand, and misapply those tools. This phenomenon is what I've called *The Owl of Minerva Problem.[4]*

Here is what I think is so interesting about the Owl of Minerva Problem, generally. Problems of intellectual life are not just those from outside the enterprise impinging upon it, but they also come from within. It's a metaphilosophical problem, because even with philosophical progress, we make errors with our best ideas. The most regular error is one I call the *hammers-and-nails* problem with philosophical insight. When you've been given a good hammer, everything starts looking like a nail. And this is definitely the case with good philosophical insights – they are deep, but we mistake that depth for universality. So, for example, pragmatists reconstruct every philosophical problem into a practical deliberation, phenomenologists reduce theoretical conflict to incompatible life experiences, and Aristotelians find a way to say that they are saving appearances (despite the fact that to those who disagree, it seems they are more stipulating them). The problem with the hammer-and-nails phenomenon is that *not everything is a nail*. In fact, *that*

---

[4] See (Aikin, 2020; Aikin and Talisse, 2020) for focused statements of the Owl of Minerva Problem.

there's a hammer at all should be evidence of that fact. Metaphilosophy, then, as an explanatory tool for philosophical dispute, and even a clarifying approach to why the debates work the way they do, occasions the disputes themselves simply climbing up to the second-order to the metaphilosophy. And so, what had promised resolution or paths forward has become a site for now-leveled up contestation (Aikin, 2023).

With argumentation, this is exactly what has happened. We learned a set of norms of reason-exchange for the sake of improving our first-order deliberations, but because we are fallible, we've developed habits of using those rules in partisan fashion, interpreting others invoking the rules to be doing so in partisan fashion, or we simply err in identifying the cases, rules, and their details in their particular complex forms.

This phenomenon takes two motivational forms, tracking the familiar distinction between *sophism* and *paralogism.*[5] To start, a *sophism* is a bad argument, given for the sake of duping an audience. So, sophisms are akin to lies, as they are intended deceptions, but instead of being deceptions about the truth of a claim, they are deceptions about reason-quality. A sophism is willful argumentative deceit. Alternately, a *paralogism* is an error of reasoning made not with an attitude about deception, but it is made with an intention to get things right and reason well. The reasoner just is in error in the process of the inference or in the procedure of its demonstration.

Returning to the analogy with sports, the sophistical version of the Owl of Minerva problem is analogous to that of *gamesmanship.* Play that is considered gamesmanship is play that uses the rules of the game in a way that creates a non-game-achieved advantage, contrary to the spirit of the game. So, in soccer the *diving phenomenon* is that of simulating being fouled and crying foul for the sake of advantage that comes from whatever sanction coming with being assessed for the foul. Diving is performed to yield a penalty kick, and so fallacy-attribution is performed for an analogous advantage. The advantage, as I can see it, is a form of meta-argumentative inference bearing on the opponent's positive case and their case against the speaker.[6] Consequently, the meta-argumentative inference comes in undercutting and vindicating forms. The undercutting form is:

---

[5] I owe Andrew Aberdein for this distinction on the fallacy fallacy, see his essay on the fallacy fallacy and connected Owl of Minerva Problem (Aberdein, 2023). For a deep dive on the distinction, see (Stevens, 2024).

[6] See Beth Innocenti's work on 'crying foul' for an analysis of the rule-invocation and enforcement in normative practices, and see her later work on paying penalties for argumentative errors (Innocenti, 2011, 2022). For accounts of meta-argumentative penalties, see the work of (Krabbe, 2003; Innocenti, 2022; Goodwin, Innocenti and Eckstein, forthcoming).

> Speaker S's argument for P is a fallacy.
> Were S to have better arguments, S would not resort to fallacious arguments.
> So, S's overall case for P is weak.

In this case, the meta-argumentative argument is at its core a speaker-regarding inference – in this case, one that runs that a speaker's given reasons are indications of the speaker's judgment of the overall set of reasons in favor of their view and that bear on the issue more generally. One should give one's best and most-compelling reasons so as to be maximally informative and efficient. We shouldn't waste our time with weaker arguments when there are stronger reasons to be given. Once we see the background features of these speaker-regarding inferences, the critical face of those reasons should emerge – that is, not only are we obliged to give our best reasons for our views, but we should give the most trenchant criticisms of our opponents. If the best one has are fallacies (for or against), it not only is telling about one's positive views but it is devastating for one's critical approach. So a vindicating form:

> Speaker S's arguments against our views are fallacious.
> Were S to have better arguments against our views, they would have given them.
> If S doesn't have a good case against our view, our views are vindicated.
> So, we have reason to hold that our views are vindicated.

The vindicating case is one that addresses not only the given arguments but the counterfactually possibly given arguments. With the fallacy fallacy (and fallacy-attribution more generally) then, one not only makes an undercutting argument against one's opponent's case, but one vindicates one's own. It is the argumentative equivalent of the dive, hoping to be given the opportunity to take a penalty kick as punishment. In this case, it's claiming vindication on the other side of a critique of the opposition's case. Such are the objectives of the sophistical version of the Owl of Minerva – one *cries foul* (please get the pun!) in order for the meta-argumentative penalty to be assessed.

The Owl of Minerva paralogism form is not a piece of adversarial strategy in the argumentative exchange, but is a kind of illusion that arises when we view our practices of disputation from the second-order. Another analogy might be useful here to display the core trouble. There is a phenomenon of grammatical over-correction with pronoun use with English speakers. The initial common error is, when speaking of oneself and another party in the nominative, to use 'me' instead of 'I' in the conjunctive form. So, one might

say "Isla and me are going to the mall." Whereas, the grammatically correct locution would be "Isla and I are going to the mall." The second-stage over-correction phenomenon is taken from seeing that 'I' is the correct usage and defaulting to it *überhaupt.* So, in the direct object, "He saw Alice and I," and in the indirect object, "He gave a gift to Alice and I," when 'me' is the correct grammatical declension of the pronoun in those cases. One becomes wary of a particular error and wary of being seen to make the error, and in correcting to avoid making it, one makes other errors. This is a pathological loop, one that ironically comes from erroneously trying to follow the rules.

Another analogy might be useful. A regular phenomenon on college campuses is that students taking abnormal psychology classes start diagnosing their fellow students with mental disorders of various kinds. The students, because they see the value of the diagnostic tools, work to become sensitive to the cases they detect, but in so doing, they over-diagnose.

With the tools of informal logic, we get a relevantly similar phenomenon. Once we learn about the fallacy of asserting the consequent, inferences to the best explanation start looking suspicious. Here's one from my life, in fact:

> A: Look, if you're driving too fast for conditions, you'll spin out and crash on this turn. Your spin and crash on this turn mean you were going too fast.
> B: Ah! But that's like inferring that something is a triangle from the fact that it's a polygon!

A is making an inductive explanatory inference, but B interprets the reasoning as deductive and sees it as an invalid instance of asserting the consequent. A little training in logic made B, ironically, insensitive to good arguments, because B was aware of the common fallacy form and over-corrected for it.

This phenomenon of over-correction is magnified by partisanship, as one is simply more (and over-) sensitive to argumentative missteps by one's interlocutors. Especially if you disagree. It isn't gamesmanship, but a special kind of motivated reasoning that's behind the fact that in sports, partisan fans of competing teams can watch the same game, but they see more fouls committed by the opposing team (Hastorf and Cantril, 1954). The same goes for fallacies – you'll hear their *ad hominems* as fallacies more often than you'll hear your own.

Importantly, this feature of motivated sensitivity to fallacies is baked into the nature of argument, as argument is intrinsically adversarial.[7] We must

---

[7] For cases for the adversariality thesis, see: (Aikin, 2011; Henning, 2018; Casey, 2020; Aikin, 2021a; Alsip Vollbrecht, 2022).

resolve disagreements, answer objections, address questions, and make progress in defending and articulating our views against a backdrop of contrasting alternatives and disagreeing (or at least non-agreeing) others. Argumentation, as a social navigation of those competing options and viewpoints will have an intrinsic element of *versus* to it, and so participating in the practice makes one ripe for this heightened sensitivity to errors. Training on the fallacies sharpens that already heightened sensitivity.[8] This, on the one hand, is good news, as we can think of critical thinking as a kind of *intellectual self-defense.* But it is also too easily something that can create an illusion of rationality around one's own views – it's harder to see one's own errors in reasoning, since one has done the reasoning. It's easy, on the other hand, to see errors in others, especially when you disagree. All this is supercharged when one's learned the language of critical thinking, because one now has terms for the errors one's opponents have made and terms of defense for one's obviously excellent views.[9]

3.

Two replies to the Owl of Minerva Problem have recently been posed by David Godden (2022) and Andrew Aberdein (2023). Godden's view of the problem is that what I've termed the sophistical version of the problem is "recalcitrant," because the participants have cynically kept "two books" of commitments – private reasons and public rationalizations. The public rationalizations are posed as sophisms, arguments the arguers do not believe are good but would be effective on target audiences. This is a problem with motivations, Godden reasons, so it is "resistant to the normal repertoire of reparative (meta-) maneuvers" (2022, 36). In this Godden and I agree – with conflicted motives, a weaponized meta-language is going to be a site for further adversarial escalation, and it will also be a source of further argumentative obfuscation. And crying foul, at the meta-level this time, will only be interpreted as further cynical performative contributions.

However, Godden holds that what I've here called the paralogistic version of the Owl of Minerva Problem is "remediable," because this is merely a matter of bringing clarity to the practice. Thus, Godden has, to one form of the problem of the Owl of Minerva, an optimistic reply that he calls the "Lark of Arete." If these simply are errors, Godden reasons, they are fixable by discursive repair. He argues:

---

[8] Critical responses to the adversariality of argument can be found in: (Cohen, 1995; Rooney, 2010; Hundleby, 2013; Stevens and Cohen, 2019, 2021; Howes and Hundleby, 2021).

[9] One worry about this trajectory of thought has been noted by Lucy Alsip Vollbrecht that it yields a form of 'intellectual grandstanding' (Alsip Vollbrecht, forthcoming).

> [M]ore optimistically … even if our meta-argumentative vocabularies
> are generated only retrospectively, they can be used prospectively
> (2022, 37)

Godden holds that though this is still an "aspirational" program, it shows that the pathological loops of meta-argumentation are not inevitable. And they can be mitigated with the right interventions. So, at least, on Godden's take, one form of the Owl of Minerva Problem is avoidable or at least manageable.

Alternately, Andrew Aberdein concedes that the Owl of Minerva Problem is a genuine worry, noting that understanding a norm makes new norm violations possible, and the concept of fallacy makes the fallacy fallacy possible. However, Aberdein argues that a virtue theory of argument offers a solution to the problem, because the problem at its core is one motivated by a base worry about argumentative vice. Aberdein's case proceeds from a structural similarity between the fallacy fallacy and moralistic vice-charging. The phenomenon of vice-charging is that of accusing a party in a dialogue of some vice of thought or morality (Kidd, 2016; Aberdein, 2023). A familiar problem with this act is that it is too easy to make the charge but not so easy to live up to the evidential burdens of substantiating those charges. So, one may accuse others of being incurious, unscrupulous, duplicitous, or credulous more easily than showing they are (and those accusations are easily inflated by the fundamental attribution error of taking situational behavior to bespeak deep features of character). Aberdein calls this meta-vice *moralism*, and he notes, "critically, as with other offspring of the Owl of Minerva, it depends on a reflective awareness of our own normative practices and it works to undermine them" (2023, 277). Moralism, on Aberdein's view, is an inflated sense of the extent to which (one's own) moral criticism is appropriate, and it is derived precisely from the fact that one has the concepts of moral evaluation, one sees their value, one engages in those evaluations, but to a degree (and with at tilted perspective) that runs contrary to the objectives of the practice. Aberdein draws the parallel with the fallacy fallacy:

> So, where the fallacy fallacy is a fallacy that arises from misapplication
> of fallacy-charging, moralism is a vice that arises from misapplication
> of vice-charging.... Moralism is not among the argumentative vices that
> I have hitherto proposed but it is clearly related to them: misplaced zeal,
> a subtype of undue willingness to engage in argumentation, and an
> unfairness to others in evaluating their arguments, a subtype of
> unwillingness to listen to others.... So I tentatively ascribe the fallacy to
> the conjunction of those two vices (2023, 277)

As Aberdein sees it, the reflective work that makes positive and pathological looping possible also makes a correction to pathological looping possible. It "enables the well-intentioned to develop their argumentative virtues" (277). If the pathological looping we see with moralism and the fallacy fallacy is a consequence of vicious insensitivity to particulars of the cases at issue, then identifying virtues of that sensitivity is the solution. But it requires a broader set of skills and positive dispositions:

> This sensitivity to circumstance requires help from other virtues, either as part of a web of interacting virtues or by appeal to regulatory virtues. With respect to argumentation, there are several candidates for the latter role, including humility (2023, 277)

In this turn, Aberdein agrees with Godden's aspirationalist turn on the solution, concluding that the fallacy fallacy and the broader Owl of Minerva problem of pathological looping is a "symptom of how reflection on our own norms can be abused but such reflection is also integral to the acquisition of virtuous character necessary for good argument" (2023, 278).

Both Godden's and Aberdein's critical responses to the Owl of Minerva Problem and the fallacy fallacy are of the *let's emphasize the positive* variety of approaches to problems of mixed bags. The Owl of Minerva problem is that reflection on our practices is not an unalloyed good, as there are both positive loops of practice-to-reflection-back-to-practice and negative. The Lark of Arete and the virtue response both are cases of pointing to the fact that there are, in fact, positive loops. An alloyed good is still a good, right?

Godden's Lark of Arete seems, further, to be a plan that runs that since we've seen the pathological looping with the Owl of Minerva Problem, we (just as we had second-order solutions to first-order problems) have further-ordered solutions to the second-order problems. And so, we in fact now have the notion of the fallacy fallacy, which we may charge when we see it. Yes, we've got a pathological loop, Godden agrees, but we've also identified the error. That's one step backward, but also a step forward. And further, we have names for those other meta-argumentative fallacies in easy parlance, too: that's why *bothsiderism, whataboutism,* and the *free speech fallacy* are theorized at all. Yes, there are meta-argumentative fallacies, but we reflective creatures have reflected upon those reflective errors and have reflective correctives.

First, I do not share Godden's optimism that reason can run ahead so felicitously with these matters. The lesson of fallacy theory is that we make errors, they are persistently attractive, and we correct them on reflection. And I expect that whatever interventions we would plan would have little uptake

until we were awash in the error. For example, the observations about the possibility of whataboutism would strike most as informal logicians quibbling, unless we were already deluged in the diversionary tactic in practice. But there is a bigger problem here – the looping continues after the corrective interventions. So, let's just take the fallacy fallacy and the charges of having committed it. Now that we have that leveled-up charge of leveled-up fallacy, we may commit yet another further leveled-up error: *the fallacy fallacy fallacy*. One commits the fallacy fallacy fallacy when one has erroneously held that the fallacy fallacy charge undercuts the inference from fallacious argument identification to the rejection of the thesis argued for. Here's a case:

> A is the proponent in an asymmetric critical dialogue with B.[10] So, A has the burden of proof solely. A's entire case depends on an argument that is demonstrably fallacious, and B identifies this and concludes that A's conclusion is unacceptable and should be rejected. A argues that B has committed the fallacy fallacy, because it does not follow that a conclusion must be rejected if the case for it is fallacious. The trouble is that failed arguments in asymmetric contexts default to rejecting the conclusion with the burden of proof. So, A's conclusion, given that they've failed their argumentative task in an asymmetric context, not only may not hold their view true, but the asymmetry allows the participants to proceed as though it is not true. A, then, in charging B with the fallacy fallacy, given the details of the argumentative exchange, has committed the fallacy fallacy fallacy.

The looping continues, and given our fallibility even when we are trying to get it right and have excellent tools, we err. Those errors are certainly more intelligible against a backdrop of partisan argument, but the motives aren't necessarily of sophistical argumentative cheating, but rather they are features of how interest and investment shade our interpretations and interventions. And especially when we must place ourselves amidst our critical exchanges, this paralogistic looping problem is baked into these practices. Again, Godden and I agree on the aspirational notes here – learning logic (and so many of our other axiological disciplines) inspires us to be better than ourselves. And we

---

[10] An asymmetric critical dialogue is one wherein only one side has the burden of proof and there is a default conclusion. So, legal contexts wherein the presumption of innocence are exemplary, but most contexts where false positives are overwhelmingly costly and false negatives are not are asymmetric in the relevant sense. So, you should, when rock climbing, assume you're not tied in, unless you've definitively checked. When handling firearms, assume they are loaded unless you're sure they aren't. And when managing matters of common sense, it's best policy to default to them unless we have good reasons to think otherwise.

agree that it is common for us to follow these routes and truly improve. But Godden and I disagree on whether we can be confident that higher-order enforcements will be effective or intelligible prior to or in the midst of controversy. We can do it *afterwards,* but that's back to the problem of the Owl of Minvera flying *at dusk,* at the end of the day after everything's been done. And it seems that even with further leveled-up correctives, new leveled-up errors emerge. It just seems to me that one can still, amidst aspiring, also be skeptical of how well it will turn out.

I believe my disagreement with Aberdein is more substantial. Aberdein holds that his virtue proposal is a "remedy" for the problem of the fallacy fallacy and the Owl of Minerva (2023, 269). The core analogy, again, is between the fallacy fallacy and moralistic vice-charging, as their etiology is similarly based on the joint vices of over-enthusiasm for correction and unfairness to others. Aberdein agrees that the fallacy fallacy arises because of the progress in a person's maturity and development as an arguer, but it is because this progress is incomplete. "A *perfectly* virtuous arguer would, ex hypothesi, only ever advance good arguments" (2023, 273). This, of course, is entirely an idealization, since at best we are "*typically* virtuous arguers," who argue well often enough but also argue badly from time to time. In those latter cases, we are not arguing virtuously. So, on Aberdein's approach, the virtue of the arguer is the primary conceptual issue. The central question of argument-evaluation is whether the arguers in question "are arguing as virtuous arguers would argue" (2023, 278). Good arguing is just how a good arguer would argue were they in the argument. Aberdein's virtue theory of argument, then, yields a very simple solution to the Owl of Minerva problem, because, by hypothesis, virtuous arguers would not commit (or fall for) these kinds of errors.

Again, because we have good evidence that the loops continue looping, we have reason to think those perfectly virtuous arguers are not possible, much less probable. Assuming we must transition from our imperfect and often vicious selves to those more perfect, the closing or correcting of one pathological loop with the virtue-inducing interventions yields the opportunity for new pathological loops. And the important thing about these loops as they begin to become more layered with reflection and axiological savvy is not just how much theoretically more challenging they are to identify and explain, but how much more dialectically difficult they will be to correct. One may be able to see how one has, say, fallaciously used an *ad hominem* abusive in inferring someone was wrong about a Shakespeare sonnet interpretation on the basis of their silly haircut, but it seems considerably less plausible that one will so easily accede to charges that one's inclusion in a debate is bothsiderist reasoning or that one's challenge of consistency of critical scrutiny is really

vicious whataboutism. Because we are not wise, the tools of wisdom will not only be misused, but they will become (because we are not wise) positive hindrances to further developments of our wisdom. I see it all as a kind of rational tragedy, really. In this Aberdein and I agree on the etiology, but we disagree that virtues are an answer, because progress toward them also brings the tools for exacerbating the problem they purportedly solve. Our goal recedes even as we progress toward it, because our progress complicates us in ways that retard our reaching the goal with the tools we currently have.

I will note further that I'm generally skeptical of the strong virtue program Aberdein's approach depends on, which analyzes argumentative quality in terms of the virtues yielding it. I think it gets the conceptual grounds backwards (as we should think of arguer virtue as a function of what quality of arguments they give, not *vice versa*). Here's one way to see the problem in high relief: returning to the fallacy fallacy. Given that we all agree that the inferential version of the fallacy fallacy is itself a fallacy, I think we can all agree that the following meta-argument (that I'll call ARG*) is a good argument in reply to someone who's committed the inferential version of the fallacy fallacy to a fallacious first-order argument (which I will call ARG). Here it is:

A: <gives ARG>. Therefore, P.
B: Ah! ARG is a fallacy, so P is false.
A <giving ARG*>: Look, B, the fact that my argument (ARG) for proposition P is fallacious doesn't mean that P is false. That's the *fallacy fallacy.*

My point here is that though ARG* is a good argument, were A a perfectly virtuous arguer, A would not be able to give ARG*, because a perfectly virtuous arguer could not have given ARG, which by hypothesis is fallacious. Recall again, that on the strong virtue theory, the primary conceptual issue for argument quality is whether the arguers in question "are arguing as virtuous arguers would argue" (2023, 278). The strong virtue theory (that explains the relatively thin concept of argument quality by the thick concepts of virtue), then, suffers from a conditional fallacy, as these posited virtuous arguers can't perfectly virtuously assess their failures or reason virtuously from their errors. Because, by the stipulation, they don't fail or have errors – that's what the *typically* virtuous do. But then we are back to what sorts the virtuous arguers and arguments from the vicious moments of that arguer's performance – is it in the argument or in the arguer? In the case of ARG*, it can't be in the virtue of the arguer, because a (perfectly) virtuous arguer wouldn't have given ARG. The virtue theory (at least as stated in terms of treating argument quality as a

matter of what a virtuous arguer would do) must be hopelessly incomplete, from this theoretical perspective, and I think that even from the participant perspective of a virtuous arguer, it will be so. It seems to me that a virtuous arguer will, perhaps, ask what a fair-minded person would do in a heated disagreement, but then their attention will be directed to the strength or weakness of the cases at hand, independent of the evaluator's virtues or those of the arguers presenting them. In a way, that's what anonymous review's all about, right? We look at the arguments independently of the arguer. Consequently, I am skeptical of the virtue theory providing a resolution to the Owl of Minerva problem, because it seems to me that the problem lies precisely in the spot virtue theories cannot theorize – where the virtues aren't or are incomplete. It's like a sunbeam being surprised there are shadows. Of course, the perfectly virtuous can't see this problem – it's not theirs. But it is ours.

4.

So far, my case has been based on the thought that our critical vocabulary is a site for either intentional abuse or escalating error, sophistical or paralogistic pathological looping. I have not argued, I should note, that these dark thoughts are reasons not to engage in theory or take up the interventions, but that because we are changed by those theories and interventions, we should expect the need for new reflexive theory and interventions as we proceed. I'm an idealist still, but like any idealist who's seen well-planned ventures fail or has been disappointed by plenty of the real, I think how we pursue the ideal needs some nuance and tempered expectations.

One place where idealism about critical thinking and argumentation theory has collided with the real is with *deep disagreement*, conflicts wherein it appears that the divergent parties do not share enough overlap of commitment or procedure to resolve their disagreement with argument. The disagreements and their arguments are interminable, and they seem to get nowhere. The first disappointment, then, is that it seems that there might be significant bounds to the tools of critical thinking – some disagreements may be deep in a way that our norms and tools for argument construction and evaluation are at their limit. Robert Fogelin held that argument is not even possible in these circumstances (Fogelin, 1985). In its stead, he proposes "persuasion" as the path to managing the disparity of views, and many have developed lies of theory around the idea that non-argumentative approaches to deep disagreements are the only plausible plans for success. In these cases, the matter is in terms of how we manage educational resources and priorities, use social pressure, or refine linguistic norms (Godden and Brenner, 2010; Mazilu, 2010; Godden, 2019). Alternately, Michael Lynch (2010) holds that we can come to the issue of the

conflict of principles orthogonally, from the perspective of a deliberative veil of ignorance – where we propose that were someone to not know beforehand what truths one were to possess and what evidence one would have access to, what resources for resolution one would prefer? Lynch's expectation is that most would prefer intersubjectively shareable sources and publicly available reasons. Finally, there are those who hold that deep disagreement is a gradable concept (as depth is a gradable concept), and so deep disagreements are matters of variant significant dialectical distances to be covered for parties to reach resolution (Aikin, 2021b; Lavorerio, 2021; Popa, 2022). In short, how many arguments must be successful in practice for resolution to be possible? The higher the number (and so, greater and wider-ranging the disparity of views and procedures), the deeper the disagreement.

The problem, as I see it, is that the notion of deep disagreement is not merely a theorist's tool, but it is also thereby a practitioner's instrument for evaluation and explanation. So, just as we have seen with the argument form of the fallacy fallacy (where one over-attributes fallacies to acceptable reasoning), we have errors in the form of what I think is a *deep disagreement fallacy*. Take a disagreement that has been longstanding, perhaps the abortion issue, as we see in Fogelin's case. Notice that once we take the issue as a deep disagreement, especially given Fogelin's take on what deep disagreements are, much of the argumentation in the dispute changes hue. Fogelin's view is that the framework propositions (about which one does not reason) are at variance on whether fetuses are people (and so morally considerable). Now, once explained in this way, argumentative exchange seems hopeless, and only non-argumentative persuasion is appropriate. There are some problems emerging once we take this approach. To start, politically, Fogelin's approach a pretty dark view of how to interact with rational others who are political equals.[11] One uses "persuasion" in a non-rational sense withthem, and to clarify what he means by that term, Fogelin approvingly quotes Wittgenstein, saying that it is like what happens "when missionaries convert natives" (1985, 6). Moreover, on Fogelin's approach there's not just a view about what's appropriate to do, but there's a prediction about what's possible. It's here that I think that Fogelin's take has a few empirical problems. To start, the debates have nevertheless evolved, as responses to reasons given in the exchanges. Judith Thomson's violinist case no longer has the abortion rights argument deny moral status to the fetus (Thomson, 1971), and Don Marquis's argument from futures like ours does not require attributing current personhood or ensoulment to fetuses (Marquis, 2007). So, the explanation of the

---

[11] Along with Robert Talisse, I've argued elsewhere (Aikin and Talisse, 2020) that this is a deeply undemocratic way to approach dispute resolution.

disagreement being rooted in those framework propositions has been shown to be incorrect, and the prediction about no further development has been falsified. Yes, the arguments persist, but they take different forms, in manners responsive to the dialectical state of play. Further, given that the arguments often track religious beliefs and interpretations of religious texts, challenges to readings of Biblical passages proposed as supporting fetal personhood have emerged (Rachels, 2003, 59). The point is that there have been continued arguments, and they have made dialectical progress. And notice that were we to have taken Fogelin's explanation and consequent recommendation seriously, we would have held it all to be impossible, and we would have treated our fellow citizens and rational interlocutors as items for manipulation. This seems like a significant moral error and important missed intellectual opportunity, and it's due to the use of concept we have of argumentative evaluation and explanation.[12]

I believe the same goes for the Lynch approach to deep disagreements – the veil of ignorance is fine for people who value bringing others around to having the truth in the first place. But how might those who reject that objective take the thought experiment? Consider the thought that were you to have the truth, and having it and living in it is a profound obligation, and there are unique and non-public means to possess and understand it, then you'd have particular obligations bearing on preserving those resources and insights. If that's all the case, then all those liberal public approaches amount to apostasy. This can be the case for the practicing committed religious lifestyles or those who hold that particular social identities afford non-publicly shareable commitments. In both cases, the call of reaching the rationally distant other on the other side of the divide is faint compared to the obligations of one's profound and normative truth. In fact, given the depth of the disagreement, it's likely that one might even hold the former conjunctive objective of reaching those others and curtailing one's overriding reasons in contempt. In this case, then, diagnosing disagreement as deep and approaching it as one in need of remediation activates recalcitrance, because it asks participants to change what reasons they give for (meta-) reasons having nothing to do with their truth. As a consequence, identifying a disagreement as deep and working to remediate it actually exacerbates the already divided circumstance.

The concept of deep disagreement, when applied to an ongoing disagreement, even if it is an accurate assessment, holds the seeds for making the disagreement even less amenable to resolution. The diagnostic and

---

[12] In my (Aikin, 2019), I argue that these concerns about Fogelin-approaches bear on political arguments, and I have argued that the democratic stakes of missed argumentative exchange are very high.

remediation tools are instruments by which, if they are communicated, the disagreements are made less plausibly resolvable.

So, just as with the fallacy fallacy, the deep disagreement fallacy comes in a few forms. The first is the hasty-assessment of a disagreement to be deep from the simple difficulty of the critical conversation. We with this instance of *specious attributions* of deep disagreement, over-attribute depth, and so miss the opportunities to address the disagreements with shared reasons. The second is the error of taking it that even if the disagreement is deep (and this assessment is correct) to reason that it must be irresolvable on the merits of the case, so one must find other paths for persuasion (so, persuasion, the veil of ignorance, negotiation). But this approach alienates all the participants from the reasons they give or are given them. This is the *inferential version* of the deep disagreement fallacy, as one infers hastily from the depth of the disagreement to the long-term hopelessness of argument in the circumstances. Finally, one may take the argument to be deep, and so then proceed to interpret all further arguments from one's opponents in light of this assessment – that they are all hopeless "arguments" and one has no real obligation to be troubled by them. Thus, a *disagreement* version of the deep disagreement fallacy. And so, like the disagreement version of the fallacy fallacy, the concept of deep disagreement colors how one interprets the arguments given in the exchanges with deeply disagreeing others, and if one takes the pessimistic view, those arguments (regardless of their cogency) will appear as failures. So, the concept of deep disagreement can make those who possess it insensitive to argument quality.

One example of the distortion that this disagreement form of the deep disagreement fallacy yields is that of explaining away purported reasons given by deeply disagreeing others as results of profound illusion or effective brainwashing. So, social conservatives argue that universities are indoctrination facilities and that liberal approaches to social issues are better explained by mental disorders than reasoning (Savage, 2005). Further, these attitudes are then bolstered by interpreting more forceful reasoning interventions from progressives as bullying or simple emoting. See, for example, in an American context, the conservative political commentator Sean Hannity replies to liberal challenges to Republican policy:

> [T]his isn't a matter of Democrats trying to promote honest debate on the issues. It's a simple matter of bad faith (Hannity, 2005, 233)

Contemporary conservatives note that liberal critique of President Trump is best explained by 'Trump Derangement Syndrome' (Campanile, 2025), and

earlier, the same was said of critics of President Bush's policies – that they were driven by a 'Bush Derangement Syndrome':

> The acute onset of paranoia in otherwise normal people in reaction to the policies, the presidency – nay – the very existence of George W. Bush (Krauthammer, 2003).

And American conservative commentator Ben Shapiro sees critical reactions on gender issues in American politics through an anti-rational lens:

> [T]hat's how the left works its magic. Feminists bully both men and women who disagree with them, while simultaneously claiming to be victimized by the patriarchal structure (Shapiro, 2013, 163)

The point here is not to identify how polarized we are (in the United States or elsewhere) but to point out how when we point out how polarized we are, we arrange inferences and reasons in a way to attribute disparate beliefs to non-rational causes and to interpret critical exchange as non-rational, too. American conservative political commentator Steven Hayward warns in *The Federalist* that "'Polarization' is Liberal Code for 'Just Surrender Already'" (Hayward, 2015). The lesson is that tools for explaining why arguments can fail have become tools for interpreting arguments, often as they are unfolding, as failing.

I've given some pragmatic arguments against two of the three leading theories of deep disagreement – the framework proposition approach and the principles and veil of ignorance approach. I've argued that they both have unique Owl of Minerva problems that emerge as what I've called the *deep disagreement fallacies* as we believe them and proceed in light of holding those conceptions of depth of disagreement. I do not believe the dialectical depth approaches to deep disagreement have the same kind of problems that are consequent of these approaches. This is because the dialectical depth approach takes deep disagreements to be different only in degree from normal disagreements, while the other approaches take deep disagreements to be different in kind. And since they are different in kind, they require different treatment. It's in this variance that the Owl of Minerva problem for deep disagreement resides. On the dialectical depth approach, however, deep disagreements are only more complex, more heated, and more wide-ranging versions of our normal disagreements, so they require only more time, patience, and attention than normal disagreements.

This is not to say that the dialectical depth approach to deep disagreement will be free from Owl of Minerva problems, but only that the concept of deep disagreement operative in this approach does not bring with it unique new problems, as it should be clear that the framework approach and principles and veil approach both do. There are, however, worries about arguments continuing in contexts of profound depth. Chris Campolo notes that under these conditions of depth, when it's not clear what the norms are or what the relevant evidence is, one can develop bad argumentative habits. We may reason badly together, but luckily reach agreement that works, and so retroactively endorse those bad reasons. Or we may agree only verbally and create a "false solidarity." Or, worse, as Campolo frames it:

> If I enter into poorly grounded reasoning with you, I may become vulnerable to your strategic sophistry (Campolo, 2005, 49)

I agree that these are all possible (and sometimes even likely) results of arguing under conditions of deep disagreement, but it is not because of the fact of deep disagreement that we have these problems. Rather, these are all unhappy possibilities of arguing in the first place. Of course, we can argue badly and luckily have fortuitous agreement, which then creates an illusion of our excellence. Of course we can have mere verbal agreement. Of course we can be exposed to manipulative rhetoric. That's why argument generally requires lots of reflective and corrective work. Maybe deep disagreement is a place where this happens more often, or as the disagreement gets deeper it becomes more likely. But I don't see these as reasons not to argue, but as reasons to be more careful as the disagreement seems deeper. These are reasons, simply, to approach complex issues with care and circumspection. And further, these are not instances of the Owl of Minerva problem for concepts of deep disagreement – these consequences do not arise from our use of the concept amidst the disagreements, but are supposed to be problems arising from the depth of the disagreement. In fact, I'm inclined to say that Campolo's case for argument-avoidance in deep disagreement has a similar pragmatic problem to Fogelin's approach – we may over-attribute the evaluations of depth and so mis-diagnose the disagreements, and we would thereby make them harder to solve than before.[13]

The result, as I take it, is that our concepts of argument and its details have potential for negative loops, the Owl of Minerva problem. This is the case with the fallacies and with deep disagreements. In the case of deep disagreement,

---

[13] For other pragmatic arguments from false positives with diagnoses of deep disagreement, see (Adams, 2005; Melchior, 2023).

particular conceptions of depth (those of clashing framework propositions and those of principles and the veil of ignorance) magnify the phenomenon, while others (the dialectical depth conception) do not. This is, I think, a good pragmatic reason to favor the dialectical depth conception of deep disagreement.

5.

My view is that our normative theories can be clarified and improved with reflection and reasoned intervention. But those practices of reflection and intervention, if they are to be effective to a degree that will be worthwhile, must not only bear clarifying visions of the norms of the practices, but they must also keep up with the ways those clarifications complicate our practices. This is particularly the case in practices where we, as participants, enforce the rules while we engage in the practice. There are ways we can mitigate some of the distortions of partisanship in those enforcements (as, for example, I think that people toward the front of the line should step in to enforce the rule of 'no cutting' in the line spots behind them, because they cannot be interpreted as having a conflict of interest), but these engagements must regularly fall to partisans with crossed purposes. The extra problem, as I've argued, has been that our tools for explaining how things can be derailed can turn into implements for new derailments. Thus, the fallacy fallacy (and the fallacy fallacy fallacy), and my proposed deep disagreement fallacy, among many others.

The aspirational edge to this story is that there is a class of argumentative error that is a product of our attempts to be rational by bringing our norms of reasoning explicitly to bear on the practice. Our reason can see paths to correction, intervention, and improvement. And we can, as our conceptions develop, try to mitigate those negative loops. This insight, I think, favors particular conceptions of deep disagreement over others. This meta-argumentative, pragmatic line of reasoning generalizes to other argumentative concepts, including those bearing on norms of free speech, procedures for engagement, and principles of interpretation. The pessimistic edge is that, as we'd seen with the initial loopings, these new orders of practice-to-reflection-to-practice loops hold opportunities for newer and harder-to-theorize and correct pathologies of meta-argumentative practice. We can manipulate some, but not all, of these practical loops. But, luckily, reflective creatures that we are, we have the opportunity to reflect after those errors to understand and

explain them. And to resolve to do better next time. That's not nothing. In fact, that's still a lot.[14]

References

Aberdein, A. (2023) 'The Fallacy Fallacy: From the Owl of Minerva to the Lark of Arete', *Argumentation*, 37(2), pp. 269–280. Available at: https://doi.org/10.1007/s10503-023-09595-9.

Adams, D.M. (2005) 'Knowing when Disagreements are Deep', *Informal Logic*, 25(1). Available at: https://doi.org/10.22329/il.v25i1.1045.

Aikin, S. (2011) 'A Defense of War and Sport Metaphors in Argument', *Philosophy & Rhetoric*, 44(3), pp. 250–272. Available at: https://doi.org/10.5325/philrhet.44.3.0250.

Aikin, S. (2019) 'Deep Disagreement, the Dark Enlightenment, and the Rhetoric of the Red Pill', *Journal of Applied Philosophy*, 36(3), pp. 420–435. Available at: https://doi.org/10.1111/japp.12331.

Aikin, S. (2020) 'The Owl of Minerva Problem', *Southwest Philosophy Review*, 36(1), pp. 13–22. Available at: https://doi.org/10.5840/swphilreview20203613.

Aikin, S. (2021a) 'Argumentative Adversariality, Contrastive Reasons, and the Winners-and-Losers Problem', *Topoi*, 40(5), pp. 837–844. Available at: https://doi.org/10.1007/s11245-020-09719-w.

Aikin, S. (2021b) 'Deep Disagreement and the Problem of the Criterion', *Topoi*, 40(5), pp. 1017–1024. Available at: https://doi.org/10.1007/s11245-018-9568-y.

Aikin, S. (2023) 'Pragmatism and Metaphilosophy', in S. Aikin and R. Talisse (eds) *The Routledge Companion to Pragmatism*. New York: Routledge, pp. 311–320.

Aikin, S. and Casey, J. (2022) *Straw Man Arguments: A Study in Fallacy Theory*. London: Bloomsbury.

Aikin, S. and Casey, J. (2024) 'The Ambitious and the Modest Meta-Argumentation Theses', *Res Philosophica*, 101(1), pp. 163–180. Available at: https://doi.org/10.5840/resphilosophica20231229114.

Aikin, S. and Casey, J. (forthcoming) 'What about whataboutism?', *Social Epistemology* [Preprint].

Aikin, S. and Talisse, R. (2020) *Political Argument in a Polarized Age*. London: Polity Press.

Aikin, S.F. and Casey, J. (2023) 'Free Speech Fallacies as Meta-Argumentative Errors', *Argumentation*, 37(2), pp. 295–305. Available at: https://doi.org/10.1007/s10503-023-09601-0.

Alsip Vollbrecht, L. (2022) 'Why We Need Skepticism in Argument: Skeptical Engagement as a Requirement for Epistemic Justice', *Argumentation*, 36(2), pp. 269–285. Available at: https://doi.org/10.1007/s10503-021-09565-z.

Alsip Vollbrecht, L. (forthcoming) 'Intellectual Grandstanding: An Epistemic Bad in Argument', *Philosophy and Rhetoric* [Preprint].

Campanile, C. (2025) 'Liberal Trump meltdown causing depression, anxiety "like 9/11," doc says', 9 February. Available at: https://nypost.com/2025/02/09/us-news/liberal-trump-meltdown-causing-depression-anxiety-like-9-11-doc-says/ (Accessed: 10 February 2025).

Campolo, C. (2005) 'Treacherous Ascents: On Seeking Common Ground for Conflict Resolution', *Informal Logic*, 25(1). Available at: https://doi.org/10.22329/il.v25i1.1043.

Casey, J. (2020) 'Adversariality and Argumentation', *Informal Logic*, 40(1), pp. 77–108. Available at: https://doi.org/10.22329/il.v40i1.5969.

Cohen, D.H. (1995) 'Argument is War...and War is Hell: Philosophy, Education, and Metaphors for Argumentation', *Informal Logic*, 17(2). Available at: https://doi.org/10.22329/il.v17i2.2406.

Cohen, D.H. (2001) 'Evaluating arguments and making meta-arguments', *Informal Logic*, 21. Available at: https://doi.org/10.22329/il.v21i2.2238.

Cotton, C. (2018) 'Argument From Fallacy', in R. Arp, S. Barbone, and M. Bruce (eds) *Bad Arguments*. Wiley, pp. 125–127.

Damer, T.E. (2012) *Attacking Faulty Reasoning*. 7th edition. Boston, MA: Cengage Learning.

Finocchiaro, M.A. (2013) *Meta-argumentation: An approach to logic and argumentation theory*. London: College Publications.

Fogelin, R. (1985) 'The Logic of Deep Disagreements', *Informal Logic*, 7(1). Available at: https://doi.org/10.22329/il.v7i1.2696.

Godden, D. (2019) 'On the Rational Resolvability of Deep Disagreement Through Meta-argumentation: A Resource Audit', *Topoi*, 38(4), pp. 725–750. Available at: https://doi.org/10.1007/s11245-019-09682-1.

Godden, D. (2022) 'Getting Out in Front of the Owl of Minerva Problem', *Argumentation*, 36(1), pp. 35–60. Available at: https://doi.org/10.1007/s10503-021-09554-2.

Godden, D. and Brenner, W.H. (2010) 'Wittgenstein and the Logic of Deep Disagreement', *Cogency: Journal of Reasoning and Argumentation*, 2, pp. 41–80.

Goodwin, J., Innocenti, B. and Eckstein, J. (forthcoming) 'The Normative Pragmatic Approach to Argumentation', in S. Aikin, J. Casey, and K. Stevens (eds) *The Routledge Handbook of Argumentation*. New York: Routledge.

Hannity, S. (2005) *Deliver Us From Evil: Defeating Terrorism, Despotism, and Liberalism*. New York: Regan Books.

Hastorf, A.H. and Cantril, H. (1954) 'They saw a game; a case study', *The Journal of Abnormal and Social Psychology*, 49(1), pp. 129–134. Available at: https://doi.org/10.1037/h0057880.

Hayward, S. (2015) *'Polarization' Is Liberal Code For 'Just Surrender Already'*, *The Federalist*. Available at: https://thefederalist.com/2015/08/07/polarization-is-liberal-code-for-just-surrender-already/ (Accessed: 14 February 2025).

Hedden, B. (2019) 'Hindsight Bias is Not a Bias', *Analysis*, 79(1), pp. 43–52. Available at: https://doi.org/10.1093/analys/any023.

Henning, T. (2018) 'Bringing Wreck':, *Symposion*, 5(2), pp. 197–211. Available at: https://doi.org/10.5840/symposion20185216.

Howes, M. and Hundleby, C. (2021) 'Adversarial Argument, Belief Change, and Vulnerability', *Topoi*, 40(5), pp. 859–872. Available at: https://doi.org/10.1007/s11245-021-09769-8.

Hundleby, C. (2013) 'Aggression, Politeness, and Abstract Adversaries', *Informal Logic*, 33(2), p. 238. Available at: https://doi.org/10.22329/il.v33i2.3895.

Innocenti, B. (2011) 'Countering Questionable Tactics by Crying Foul', *Argumentation and Advocacy*, 47(3), pp. 178–188. Available at: https://doi.org/10.1080/00028533.2011.11821745.

Innocenti, B. (2022) 'Demanding a halt to metadiscussions', *Argumentation*, 36(3), pp. 345–364. Available at: https://doi.org/10.1007/s10503-022-09569-3.

Kidd, I.J. (2016) 'Charging Others with Epistemic Vice', *The Monist*, 99(3), pp. 181–197. Available at: https://doi.org/10.1093/monist/onv035.

Krabbe, E.C.W. (2003) 'Metadialogues', in F.H. Van Eemeren et al. (eds) *Anyone Who Has a View*. Dordrecht: Springer Netherlands (Argumentation Library), pp. 83–90. Available at: https://doi.org/10.1007/978-94-007-1078-8_7.

Krauthammer, C. (2003) 'The Delusional Dean', *The Washington Post*, 4 December.

Lavorerio, V. (2021) 'The Fundamental Model of Deep Disagreements', *Metaphilosophy*, 52(3–4), pp. 416–431. Available at: https://doi.org/10.1111/meta.12500.

Linker, M. (2014) 'Epistemic privilege and expertise in the context of meta-argument', *Argumentation*, 28. Available at: https://doi.org/10.1007/s10503-013-9299-6.

Lycan, W.G. (1996) 'Consciousness and Experience', *Philosophy*, 72(282), pp. 602–604.

Lynch, M.P. (2010) 'Epistemic Circularity and Epistemic Incommensurability', in A. Haddock, A. Millar, and D. Pritchard (eds) *Social Epistemology*. Oxford University Press, pp. 262–277. Available at: https://doi.org/10.1093/acprof:oso/9780199577477.003.0013.

Marquis, D. (2007) 'An Argument That Abortion is Wrong', in R. Shafer-Landau (ed.) *Ethical Theory: An Anthology*. Wiley-Blackwell, pp. 439–450.

Mazilu, S. (2010) 'THE ABORTION DEBATE AS A CONTEST OF INCOMMENSURABLE MORALITIES', *Interstudia (Revista Centrului Interdisciplinar de Studiu al Formelor Discursive Contemporane Interstud)*, (06), pp. 220–227.

McKeon, M.W. (2021) 'Inference Claims as Assertions', *Informal Logic*, 41(3), pp. 359–390. Available at: https://doi.org/10.22329/il.v41i3.6844.

McKeon, M.W. (2022) 'Arguments and Reason-Giving', *Argumentation* [Preprint]. Available at: https://doi.org/10.1007/s10503-021-09561-3.

Melchior, G. (2023) 'Rationally Irresolvable Disagreement', *Philosophical Studies*, 180(4), pp. 1277–1304. Available at: https://doi.org/10.1007/s11098-023-01933-7.

Miller, J.S. and Miller, C.K. (2015) *Why Brilliant People Believe Nonsense: A Practical Text For Critical and Creative Thinking*. Wisdom Creek Academic.

Pinto, R.C. (2001) 'The Relation of Argument to Inference', in Pinto, R.C., *Argument, Inference and Dialectic*. Dordrecht: Springer Netherlands, pp. 32–45. Available at: http://link.springer.com/10.1007/978-94-017-0783-1_4 (Accessed: 22 February 2023).

Popa, E.O. (2022) 'On the Rational Resolution of (Deep) Disagreements', *Synthese*, 200(4), pp. 1–15. Available at: https://doi.org/10.1007/s11229-022-03753-4.

Rachels, J. (2003) *The Elements of Moral Philosophy*. 4th edn. Boston, MA: McGraw-Hill.

Rooney, P. (2010) 'Philosophy, Adversarial Argumentation, and Embattled Reason', *Informal Logic*, 30(3). Available at: https://doi.org/10.22329/il.v30i3.3032.

Savage, M. (2005) *Liberalism is a Mental Disorder: Savage Solutions*. Thomas Nelson Inc.

Shapiro, B. (2013) *Bullies: how the left's culture of fear and intimidation silences America*. New York: Threshold Editions.

Stevens, K. (2022) 'Silence at the Meta-Level: A Story About Argumentative Cruelty', *Philosophy and Rhetoric*, 55(1), pp. 76–82.

Stevens, K. (2024) 'Sophisms and Contempt for Autonomy', *Philosophy & Rhetoric*, 57(3), pp. 333–346.

Stevens, K. and Cohen, D. (2019) 'The attraction of the ideal has no traction on the real: on adversariality and roles in argument [1]', *Argumentation and Advocacy*, 55(1), pp. 1–23. Available at: https://doi.org/10.1080/10511431.2018.1504584.

Stevens, K. and Cohen, D.H. (2021) 'Angelic Devil's Advocates and the Forms of Adversariality', *Topoi*, 40(5), pp. 899–912. Available at: https://doi.org/10.1007/s11245-020-09726-x.

Thomson, J.J. (1971) 'A Defense of Abortion', *Philosophy and Public Affairs*, 1(1), pp. 47–66.

Wohlrapp, H. (2014) *The Concept of Argument: A Philosophical Foundation*. Dordrecht: Springer.

Scott Aikin
Vanderbilt University
scott.aikin@vanderbilt.edu