Fallacy Theory, the Negativity Problem, and Minimal Dialectical Adversariality

Teoría de la falacia, el problema de la negatividad, y adversarialidad dialéctica mínima

Scott F. Aikin

Vanderbilt University, Nashville, United Sates
scott.f.aikin@vanderbilt.edu

Abstract: Fallacy theory has been criticized for its contributing to unnecessary adversariality in argument. The view of minimal adversariality by Trudy Govier has received similar criticism. A dialectical modification of Govier's minimal view is offered that makes progress in replying to these challenges.

Keywords: Fallacy theory, argumentation, Adversary Paradigm.

Resumen: La teoría de la falacia ha sido criticada por su contribución a una adversarialidad innecesaria en el argumento. La perspectiva de la adversarialidad mínima de Trudy Govier ha recibido críticas similares. Una modificación dialéctica de la perspectiva de Govier se ofrece para progresar en la respuesta a esos desafíos.

Palabras clave: Teoría de la falacia, argumentación, paradigma adversativo.

1. Introduction

Fallacy theory is the convergence of three broad programs in the study of argument. First is the program of defining what fallacies are and taxonomizing their types. Second is the pedagogical program of teaching some taxonomy of fallaciesand the skills of their detection and correction as part

of critical thinking classes. Third is the meta-theoretical program of articulating what the relationship is between understanding fallacies and the broader program of understanding arguments and reasoning in general.

Fallacy theory has come under considerable criticism of late. Of particular interest and urgency has been the *negativity problem*: foregrounding failure and the vocabulary of criticism promotes argumentative adversariality, and as a consequence contributes to bad argumentative practice.

This paper is a reply to this concern about fallacy theory. The reply is that argumentative exchange is best conceived as *dialectically minimally adversarial*, and so fallacy theory must then provide tools for articulation of criticism and also the tools for management and de-escalation of critical discussion.

My plan here is to briefly survey what I see as the three domains of fallacy theory, then turn to what I take as the line of criticism identified by the negativity problem. In particular, the negativity problem is most clearly manifested in what is termed the Adversary Paradigm for argument, that critical exchanges are contests and interlocutors are opponents. This, critics reason, yields worse argumentative outcome. Finally, my modest defense of fallacy theory will be to concede much of the critical bite of the cases against fallacy theory but to hold that these are welcome occasions for reform and reconception. In particular, the thought behind Trudy Govier's notion of minimal adversariality in argument can be preserved with a qualification of reason-exchange in critical dialogue.

2. Fallacy theory and its components

Fallacy theory is a subdomain of argumentation theory. A commonplace is to contrast the focus of this broader domain with that of formal or deductive logic; the latter concerned with conditions for argumentative validity and the former concerned with the weaker forms of support for arguments as products and other procedural issues with argument as process. Fallacy theory is the more restricted study of ways support fails or procedural rules of argument are broken. Exactly how to even thematize these failures is precisely one of the core issues of fallacy theory. And so, there are divisions about how to even define what a fallacy is. There is the 'standard

treatment,' as identified by Hamblin (1970), that fallacies are arguments that seem valid but are not. There is the broadened version, as developed by Johnson (1987, p. 246), that a fallacy is an argument that violates one of the standards for good argument and occurs with sufficient frequency to merit being classified. Further, there is the pragma-dialectical perspective, as seen with van Eemeren and Grootendorst (1987, p. 297), that fallacies are discussion moves that threaten the resolution of a dispute – and in particular, they are violations of rules of critical discussion. Alternately, a fallacy may be, as Walton (1995, p. 15) terms it, the misuse of an argument scheme. There are, of course, more varieties of definition, and they generally depend on the theory of argumentative normativity on offer, as all theory of fallacy is a theory of how one *fails* to do what one *ought* in argument. Disagreement about argumentative norms yields disagreement about what it is to break those norms or fail their demand.

The second focus for fallacy theory is on how informal logic is taught in the classroom. Again, a contrast with formal logic is useful. With natural deduction, the focus is on rules of good inference and their systematicity, particularly in construction of proofs. Little systematic effort is put into the articulation of ways to fail at the objectives of proof. In contrast, the overwhelming amount of time and energy put into classroom work in and textbook space in informal logic is on fallacies – how arguments fail. And so training for students is often in the form of fallacy-spotting, not argument construction. As Johnson notes, most fallacy terminology is in the sake of "initial probing" (1987, p. 248), and so the accurate use of fallacy charges is part of training in productive dialogue. Work in fallacy theory informs pedagogy in the sense that well-taxonomized and explained accounts of fallacy allow students a rich interpretive framework for discussion.[1]

The third, metatheoretical, component of fallacy theory is the task of articulating how findings in fallacy theory inform our broader research of

---

[1] As Hundleby observes (2009), many textbooks fail to have well-developed fallacy theoretic discussions beyond presenting the standard fallacy forms (the gang of 18). This, I believe, is not the fault of fallacy theorists in the first instance, but the fault of publishers for producing textbooks with so little responsible engagement with these domains. However, fallacy theorists do bear some responsibility in the second instance, as they should be writing more textbooks showing the significance of responsible fallacy theory. Robert Talisse and I have tried to correct this imbalance with our *Why We Argue* (2014).

argumentative and discursive axiology. What does a certain fallacy reveal about argumentative norms? What does the prevalence of a class of vicious dialectical tropes tell us about our society? How does argument, even though we are regularly bad at it, fit with democracy? A natural thought is that certain argumentative failures are pregnant with meaning about what argument should be, how it should work. And so, out of a few object lessons in how not to argue, we have information about how to argue. And so, a kind of reflective equilibrium arises between our theories of argument and our systematic treatment of fallacy. That's the hope.

3. The negativity problem

Fallacy theory is a systematic view of argumentative *error*. The vocabulary of fallacies, as a consequence, is univocally *critical*. There are two consequences of this negative-emphatic view. The first is that fallacy theory has a *problem with misplaced emphasis* – we should not only be looking for ways to criticize arguments, but to construct good ones and improve bad ones. The second is that fallacy theory, in its negativity, is complicit with (and promotes) the excessive adversariality of argumentative exchange.

Extending Janice Moultin's (1983) criticism of the Adversary Paradigm in philosophy, Catherine Hundleby (2009, 2010, and 2012) and Phyllis Rooney (2010 and 2013) have argued along both of these lines. Because of fallacy theory's negative valence, negative consequences ensue. Hundleby observes that: " The oppositional nature of fallacy-allegation ... lends itself to formulations according to the politically regressive and epistemologically archaic Adversary Paradigm (2010, p. 280)" .

Hundleby further observes that the way fallacies are regularly presented in textbooks offer "no suggestion of argument repair" (2010, p. 289) and yield "pin the tail on the argument" exercises for students.[2] Phyllis Rooney, similarly, argues that the adversarial paradigm is epistemically and argumentatively stunted: "[T]he Adversary Paradigm either leads to bad reasoning ... or ... it sustains a more limited range of reasoning and argument

---

[2] One important outlier on this point is Richard Epstein and Michael Rooney's *Critical Thinking* (2017), which has exercises in argument repair.

forms ... (2010, p. 205)" . In short, the negativity of fallacy-identification is part of and contributes to the Adversarial Paradigm, which obscures the goals of truth-seeking.[3]

Further, Hundleby observes that the domain of most fallacy theory is from the perspective of those who are roughly social equals trading reasons. This, for sure, is a relevant domain, but it is not exhaustive of the scope of bad and recurrent argument types. Hundleby observes that too much is left out – there are 'androcentric fallacies' (2009: 2), and there is a growing literature on how too many from underrepresented groups are not given their due in critical dialogue, beginning with epistemic (Fricker 2007 and Medina 2013) and extending to argumentative (Bondy 2010, Rooney 2012, Hundleby 2013, and Heikes 2017) injustices. And given the adversarial model for fallacy theory, we obscure the ways social inequalities are exerting pressure on how we assess arguments.[4]

4. The Modest Defense

The modest defense of fallacy theory is to concede the negativity problem. Fallacy theory, taught exclusively, yields sharks, not arguers. It is a common phenomenon, when teaching a survey of informal logic, to have students ask whether there are any good arguments.[5] However, there are things to be said in favor of fallacy theory despite the negativity problem.

There are two parts to the *modest* reply: the mutuality thesis, and the intrinsic adversariality thesis. The mutuality thesis is that vocabularies of negative assessment are both part of normative vocabularies and important to their development. *Insofar as there is 'ought,' there are correlate*

---

[3] Others who have noted this connection are: Paul (1984), Johnson (2000), Cohen (2004) and Boudry, Paglieri, and Pigliucci (2015).

[4] The double-bind problem for women is exemplary of this trouble. The adversariality of argument puts women with particular social expectations at serious disadvantage, as they are expected to be demure, and so must forego the most vigorous exchanges. If they do take part vigorously, they pay social costs for it that men do not. (See Hundleby 2013, Rooney 2012, and Gilbert 2014).

[5] A running joke with teachers of critical thinking classes is the analogy between their students and those taking a survey of abnormal psychology. One group of students think that all arguments are to be classified as *ad hominem* or false dilemma fallacies, while the other group thinks that everyone in the dorm suffers from multiple personality disorder.

'ought-nots' that clarify and provide application. The intrinsic adversariality thesis is that a minimal degree of dialectical adversariality is part of any argumentative exchange; so as a consequence, negativity is an inescapable component of argument, and any proper theory of argument must be poised for the proper management of fallacy-vocabulary, not its elimination. I will argue first for mutuality, then for the more controversial intrinsic adversariality thesis.

5. The Mutuality Thesis

The argument for mutuality begins with what I take to be a truism about normative practices – all normative practices have metalanguages formulable about them. For example, natural languages have grammars, but the language of grammar need only be *possible*. Or take poetry. The vocabulary of rhyme schemes and metaphor need only be possibly formulated about poetry. The same, I think, goes for logic. We have reasoning and arguments, but the vocabulary of logic (formal and informal) is a metalanguage that makes the rules of the first-order practice explicit. So, the metalanguage of logic is a repository of the rules we (ought to) follow when we reason. All normative practices have the possibility of *error* in their performance, because being bound by rules doesn't guarantee that they are followed. In the case of grammar, common errors are called, for example, run-on sentences, failures of parallel, subject-verb disagreement, and so on. With poetry, we have bad or mismatched rhymes and mixed metaphors. The same goes for fallacy theory – common error types are theorized and given names. The point of many of these metalanguages is not only to make the norms explicit, but to facilitate the function of first-order normative practice. That is, with both grammar and logic, the point is to make the rules and errors explicit not just for their own sake, but for the sake of successful and reflective normative practice.

When a normative practice is self-consciously assessed, the variety of errors clarifies the norms and the newly clarified norms allow practitioners to refine their first-order practices and also to find errors that had previously escaped their critical gaze. And so with a metalanguage, particularly the metalanguage of criticism, normative practices evolve as the kinds of prac-

tices we can self-reflectively endorse. Again, writing classes are improved with the introduction of the second-order language of grammar, and so writing on the first order is improved, too. Similarly with poetry, poets who have the concept of rhyme scheme and a notion of what mixed metaphors are are more effective poets, since they can craft their poems with an eye to particular achievements and avoiding certain errors. By analogy, this reasoning bears on argument. With the concepts of *premises, conclusions*, and *inferences,* we may not only reason better, but we may more self-consciously make that reasoning manifest to one another. And so the vocabulary of criticism, with the concept of the fallacy and its particulars, such as straw figure or false dilemma, we may correct errors with a vocabulary that makes a road to epistemic improvement clearer.

Fallacy theory, then, is part of a larger dialectic of rationality unfolding, the norms of reason exfoliating against where we make errors in trying to follow and enforce them.[6] And so we see the importance of their belonging to a language of argument and argument-assessment, since they not only allow for criticism of particular arguments, but also for assessments of well-run dialectical exchange. Fallacy theory amounts to the development of a critical metalanguage that is dialectically heuristic in its first-order application, but is in the service of broader norm clarification. The process is open-ended, because the phenomenon explained and discovered is a moving target – how we argue is, in many ways, influenced not only by what we are reasoning about, but also in terms of *how we critically talk about our reasoning.*[7] The norms, then, will and ought to evolve as we develop norms of criticism. Negativity, then, is a necessary component of self-conscious reasoning. But this is not to say that the negativity should be our sole focus. In fact, the negative critical components of fallacy theory need to be integrated with other programs and objectives, such as argument repair (as noted by Hundleby 2010). However, argument repair is impossible unless we have identified and made explicit what's wrong first.

---

[6] This Hegelian thought, I think, is best captured by Robert Brandom's memorable slogan, "logic is the organ of semantic self-consciousness" (1994: xix). With logic, we make clear what we mean, how we grasp it, and who it is relevant to the other thoughts we have. We, in effect, know what we are doing when we are thinking.

[7] See Aikin and Talisse (2017) for an account of the 'Owl of Minerva Problem' that arise from this necessarily backward-looking nature of norm-clarification.

6. The Intrinsic Adversariality Thesis

The intrinsic adversariality thesis is considerably more controversial than the mutuality thesis. Both Hundleby (2009, 2010, and 2013) and Rooney (2010 and 2012) have argued that adversariality may be a dominant paradigm, but it is both a bad and an optional one. Consequently, it should be foregone. More cooperative models are available, so intrinsic adversariality is indefensible. Further, it is clear that the presence of adversariality in argumentative exchange can subvert the broader epistemic objectives of argument. Finally, adversariality puts many disadvantaged groups at further disadvantage, and it creates retrenchment in the face of further criticism.

Hundleby and Rooney's arguments target primarily Trudy Govier's case for minimal adversariality. Govier's model proceeds from the premise that if one's audience must be on the receiving end of an argument, one must presume that they need correction. As Govier frames the argument:

1. I hold X
2. I think that X is correct. (Follows from 1)
3. I think that non-X is not correct. (Follows from 2)
4. I think that those who hold not-X are wrong, or are making a mistake (Follows from 3)
5. Should I need to argue for X, I will thereby be arguing against not-X.
6. Those who hold not-X, are, with regard to the correctness of X and my argument for X, my opponents (1999, p. 244).

The objection is to line 6. As Hundleby puts it, the trouble is that "we may exchange reasons without opposing one another's ideas – never mind opposing one another personally" (2013, p. 239). Govier's core thought is that argument is part of the management of rational resolution of controversy. So "when we argue for a claim, we at the same time, and necessarily, argue against an envisaged opponent, one who does not accept the claim" (1999: 243). Govier distinguishes between this *deep, minimal, and intrinsic* adversariality and the *negative ancillary* adversariality, one may see with raised voices, dismissive gestures, and expressions of contempt. But, again, the challenge from Rooney and Hundleby is that even the deep, minimal

and intrinsic adversariality picks out an *opponent*. This yet yields a distortion, since the challenge is to represent what is *rational* in the opposition, not that it is oppositional. This objection is correct, and I think it scores the right critical challenge on Govier's model.

The modest strategy for defending fallacy theory along these lines of adversariality is to concede Hundleby and Rooney's point, but make the case for a revised notion of minimal adversariality. Call this particular version *dialectically minimal adversariality*. Here is a schematized version of the case:

1. If arguer A is arguing for a view, p, then p is either controversial or potentially so, and A is seeking to resolve that controversy.
2. If a view, p, is controversial, then there is someone who either holds that p is false or dubitable. If p is potentially controversial, there is a possible reasonable perspective that takes it that p is false or dubitable. Call any actual or potential holder of these perspectives W.
3. If 1 and 2, then W is A's target audience for the argument for p.
4. If W holds that p is false or dubitable, then W provides one or more of these critical challenges:
   (a) W has *rebutting* reasons against p (showing that p is false, or showing that some other claim, q, is true, which is a contrary of p), or
   (b) W has *undercutting reasons* against p (ones that show that p, given the evidence, is not likely true or reasonable)
5. If A takes W as the target audience, then A addresses W's critical challenges with either:
   (a) vindicating reasons – reasons that establish p in full, or
   (b) revising reasons – reasons that establish p*, which is a weaker or more defensible version of p.
6. *Therefore*, if A is arguing for p, then A is giving reasons within the following dialectical complex:

Addressing \ In form of	Vindicating reasons	Revising Reasons
To W's Undercutting challenges	V-p to U	R-p* to U
To W's Rebutting challenges	V-p to R	R-p* to R

7. All reasons within the dialectical complex are reasons answering critical challenges.
8. *Therefore*, All argument-giving is with the objective of answering critical challenges.

Call this revision of Govier's minimal adversariality view *dialectically minimal adversariality*. The only adversariality in this model is the matter of weighing the force of the better reasons, and so this is minimal and only dialectically adversarial. As a consequence, the force of this notion of dialectical adversariality is in the reasoned weighing of evidential considerations for and against a view. Moreover, given the way the program is maintained, the adversariality can be rationally managed, since with revising reasons, one can concede a point made by a critic without having to evacuate one's view. The result is that the notion of an arguer taking a critical line with one's own commitments is, for the purposes of the argument, *both* an opponent and an ally. The thought is that without the role-related duties of critical dialogue, there are moves of critical probing that must be performed that are, in their dialectical function, *oppositional*. However, this is yet in the service of a broader *cooperative* goal of dialectical testing of reasons and acceptability. Consequently, we may, within this dialectical representation, provide a model whereby we are not opposing one another personally, but can maintain the role designed duties of critics and defenders of views under scrutiny.

The upshot of this argument is that the adversariality of addressing critical challenges to a view is an essential part of argumentation. One must either speak to a critic or construct a motivating reason for marshalling one's reasons in the form of a potential challenge. Moreover, there is nothing to this program that excludes the Gricean virtues of cooperation, as it surely takes the cooperative principle to interpret the significance of challenges, objections, and critical questions.

If this line of reasoning is correct, the crucial element to training in fallacy theory is to mitigate the escalation of adversariality from the exchange of critical feedback to the exchange of insults. So long as the cooperative exchange is critical discussion, dialectically minimal adversariality need not be any impediment to arguments given in the spirit of weighing com-

peting reasons.[8] Moreover, notice that there can be collaborative elements that emerge from these considerations such that one can truly value critical questions and challenges, not as personal attacks or even rejections of one's point of view, but as the kind of useful resistance needed to craft the case for any controversial view.[9]

7. Conclusion

Fallacy theory, properly framed, is a domain with contested target phenomena, and as a consequence, contested applicability and normativity. This comes as no surprise to anyone familiar with philosophy's long history of fits and starts on a variety of issues. The phenomenon of reasoning is a moving target, as our vocabularies of evaluation change the phenomenon. In fact, the important thing is that they *do* change the phenomena. That people now invoke *ad hominem* attacks or slippery slopes in the midst of arguments is testament to the contribution fallacy theory makes to self-conscious argumentation. Further, the negativity of this critical vocabulary is itself intrinsic to this program of bringing normative practices to self-consciousness. The terms of critique are part of both the first-order practice of argument (as the minimal dialectical adversariality argument runs) and is part of our grasp of that practices as rule-bound (as the mutuality argument runs). The conclusion, then, is that fallacy theory is messy and adversarial, and necessarily so. What's required of us, then, as argumentation theorists, is not that we reject fallacy theory or reform it to the point of being non-adversarial, but that we develop research and teaching programs that (a) maintain a minimum of well-orderedness to research and (b) mitigate the potentiality of adversarial escalation in argument. So programs of *argument repair* alongside fallacy identification must be taught,

---

[8] See Aikin (2011) for models of argumentative escalation and accounts of its mitigation.

[9] Note in this regard that dialectical minimal adversarialism maintains and even highlights what Govier (1999:250) took to be her motivation for the intrinsic adversariality thesis: the Millian commitment to the thought that *knowing only one's own side of a debate is knowing little of that.*

and we must keep track of the way our critical vocabulary returns to influence the practices it is designed to describe.[10]

Works Cited

Aikin, Scott. "A Defense of War and Sport Metaphors for Argument." *Philosophy and Rhetoric*, 44 (2011): 3.

Aikin, Scott and Talisse, Robert. *Why We Argue (and How We Should)*. New York: Routledge, 2014.

Aikin, Scott and Talisse, Robert. "Democracy, Deliberation, and the Owl of Minerva Problem." *The Critique,* January/February. Available at (2017): http://www.thecritique.com/articles/democracy-and-the-owl-of-minerva/

Bondy, Patrick. "Argumentative Injustice" *Informal Logic,* 30 (3) (2010): 263-278.

Brandom, Robert. *Making it Explicit*. Cambridge: Harvard University Press, 1994.

Cohen, Daniel. *Arguments and Metaphors in Philosophy*. University Press of America, 2004.

Epstein, Richard and Rooney, Michael. *Critical Thinking; Fifth Edition*. Socorro, NM: Advanced Reasoning Forum, 2017.

Fricker, Miranda. *Epistemic Injustice*. New York: Oxford University Press, 2007.

Gilbert, Michael. *Arguing with People*. Toronto: Broadview Press, 2014.

Govier, Trudy. *The Philosophy of Argument*. Vale Press, 1999.

Hamblin, Charles. *Fallacies*. Metheuen: London, 1970.

Heikes, Deborah. "On Being Reasonably Different." *Southwestern Philosophy Review*, 33 (2017): 52-63.

Hundleby, Catherine. "Aggression, Politeness, and Abstract Adversaries." *Informal* Logic, 33 (2) (2013): 238-262.

Hundleby, Catherine (2010) "The Authority of the Fallacies Approach to Argument Evaluation." *Informal Logic*. 30:3, 279-308.

Hundleby, Catherine. "Fallacy Forward: Situating Fallacy Theory." In Juho Ri-

tola (ed.), *Argument Cultures: Proceedings of OSSA 2009*. Windsor, ON: OSSA. 1-10, 2009.

Johnson, Ralph H. "In the Blaze of Her Splendors: Suggestions about Revitalizing Fallacy Theory." Argumentation, 1 (1987): 239-253.

Johnson, Ralph H. *Manifest Rationality*. Mahwah, NJ: Lawrence Erlbaum Associates, 2000.

Medina, Jose. *The Epistemology of Resistance*. New York: Oxford University Press, 2013.

Moulton, Janice. "A Paradigm of Philosophy: The Adversary Method." In Sandra Harding and Merrill Hintikka (eds.). *Discovering Reality* (pp. 149-164) Boston: D. Reidel, 1983.

Paul, Richard. "Teaching Critical Thinking in the 'Strong' Sense." *Informal Logic Newsletter*, 4 (2) (1984): 2-7.

Rooney, Phyllis. "When Philosophical Argumentation Impedes Social and Political Progress." *Journal of Social Philosophy*, 43 (3) (2012): 317-333.

Rooney, Phyllis. "Philosophy, Adversarial Argumentation, and Embattled Reason." *Informal Logic*, 30 (3) (2010): 203-34.

Walton, Douglas. *A Pragmatic Theory of Fallacy*. Tuscaloosa: Alabama University Press, 1995.