ORIGINAL RESEARCH



On Halting Meta-argument with Para-Argument

Scott Aikin¹ • John Casey²

Accepted: 26 January 2023 / Published online: 12 February 2023 © The Author(s), under exclusive licence to Springer Nature B.V. 2023

Abstract

Recourse to meta-argument is an important feature of successful argument exchanges; it is where norms are made explicit or clarified, corrections are offered, and inferences are evaluated, among much else. Sadly, it is often an avenue for abuse, as the very virtues of meta-argument are turned against it. The question as to how to manage such abuses is a vexing one. Erik Krabbe proposed that one be levied a fine in cases of inappropriate meta-argumentative bids (2003). In a recent publication (2022) Beth Innocenti expands on this notion of a penalty, arguing that some meta-arguments should be halted with "shouting, cussing, sarcasm, name-calling." In this essay, we review Innocenti's case that these confrontations and haltings improve the argumentative circumstances. We provide three reasons that this promise is not well-founded. First, that such confrontations have a significant audience problem, in that they are more likely to be interpreted as destroying the argumentative context than improving it. Second, that Innocenti's procedural justification, that those who lose meta-discussions should pay a penalty, is not satisfied if the metadiscussion is halted. And third, there is a boundary problem for the cases, since it seems there is no principled reason to restrict halting meta-arguments just to these cases (especially if there is no meta-discussion on the matter to make the bounds explicit). Though expressions of anger can be appropriate in argument, we argue, it cannot take the place of argument.

Keywords Meta-argument · Para-argument · Fallacy-accusation · Beth Innocenti



Scott Aikin scott.aikin@vanderbilt.eduJohn Casey j-casey1@neiu.edu

Vanderbilt University, Nashville, USA

Northeastern Illinois University, Chicago, USA

1 Introduction

When we have critical discussions, we often must pause to have critical discussions about those critical discussions. So, there are arguments that occasion meta-arguments. With the arguments, we settle questions over items of first-order interest—for example, how many cats to adopt or whether Finland is a nice place to visit in the winter. With meta-arguments, we evaluate and clarify those first-order arguments—for example, which disputant has the burden of proof, whether a given argument was fallacious, and when we've reached a conclusion and can move on. Arguments are costly in terms of the energy and time one must put into them, and meta-arguments are costly too, in similar fashion and for additional reasons. In particular, they also require significant attention to detail with the given arguments, clarity with regard to the explicit (and implicit) norms of critical discussion, and they have the added opportunity costs of participants not pursuing resolution on the topic of first-order interest. Meta-discussion is, then, too often an irksome distraction.

Given that both arguments and meta-arguments are costly, there are epistometric and social considerations bearing on whether to have them at all (Paglieri 2009; Paglieri and Castelfranchi 2010; Hample et al. 2012). Is it worth the expenditure on either of these levels? Given that it is costly to have meta-arguments particularly, because they require considerable attention to what has been said and clarity on rules of argument and because it is a diversion from the main object of contention over a point of order, it has been proposed that those who do not prevail in those meta-disputes should have a penalty assessed for their bad investment of intellectual capital. Erik Krabbe holds that the participant who has not "won the metadialogue ... is to pay the costs of the metadialogue," or be "punished with a fine" (2003, 89). More recently, Beth Innocenti argues that the rule that we "penalize an interactant who has not 'won the metadialogue' as a way of halting retreats to metadialogue deserves to be taken seriously" (2022, 359).

The question of what the penalty is and how to have an interlocutor pay it looms. Argument is not a site where one can pay such fines, since if one has lost the meta-argument, then not having one's contention upheld is the extent of the argumentative payment. So, there's a sense that the fines are already internal to the dialogue—in this case, in not having one's proposal accepted. But this is not the relevant sense of *fine* or *sanction* proposed. So, it seems something, instead of internal to, *adjacent to argument* must be the site of this enacted punishment or fine. Call it *para-argument*, something alongside of, but dependent on and supplementing, argument. And the issue is how one pays for one's meta-argumentative errors and failed meta-discussions in para-argument. So, for example, the winner of the meta-argument may be allowed to speak condescendingly for a time to a losing conversant over the issue. Or one may exclaim, "Swing and a miss!" after having rebutted the accusation of fallacy. Or one may, after having shown that an argument begs the question, joke

¹ See also Hamblin (1970, 284) for the observation about 'points of order.' And for further notes on norms of meta-dialogue and criticism, see van Eemeren and Houtlosser (2009), Bermejo-Luque (2010), Innocenti (2011), and Linker (2014).



that witnessing the argument unfold was like watching a magician stuff a hat with rabbits.² These are all mild chidings, and they both reinforce the norms of argument and extract the para-argumentative payment from those who have lost the meta-argument. The fine is paid in the form of bearing some laughter-at or a dismissive tone, or even explicit reprimand. The objective being that our conversant is in the stocks for this treatment only momentarily, and we can return to the argument with the point of order in question clarified. And the purpose of the payment extracted is to repair the argument exchange and return to the matter at hand, without further distraction over further unnecessary points of order.

Innocenti argues that such para-argumentative penalties should be proportionate to the errors and harms done in them. And they should not only be done after failed meta-discussions but in advance in order to *halt some of them altogether*. She proposes methods to "get addressees to stop retreating to metadiscussions that derail ground-level discussions" (2022, 345). Innocenti further asserts that some errors and harms of those proposed meta-discussions are great enough to merit very confrontational language and behavior, including "shouting, cussing, sarcasm, (and) name-calling" (2022, 345). While she concedes that these "strategies ... may at first glance appear to be out of bounds for ideal critical discussion," given the stakes of the first-order critical discussion, it is important that the costs for opening inappropriate metadialogue be very high, so such para-argumentative prevention can be deterrents to further abuse (2022, 345).

It is useful to view Innocenti's proposal in light of recent work on argumentative adversariality. Consider Trudy Govier's famous distinction between minimal and ancillary adversariality (2000, 2021). Minimal adversariality is what follows necessarily from not sharing the same view, and ancillary adversariality is the unnecessary unpleasantness—the finger pointing, the name-calling, "shouting, cussing and sarcasm" that Innocenti is proposing here. The broad, though by no means universal, trend in recent literature in argumentation studies has been to view adversarial features of arguing with suspicion. Indeed, in some quarters, the worry is that adversariality of the *minimal* type is already *too much*, and it is to be blamed for much that is bad in argument.³ Further, shouting and so forth are extra-argumentative, purely adversarial contributions. Few have stood up for them in any context,⁴ so Innocenti's proposal has significance well beyond the conduct of meta-arguments. If adversarial sanctions serve meta-argumentative ends, then it is likely that they could play a role in the first order.

Here, we will critically review Innocenti's case for escalations in para-argumentative adversariality and her proposal that penalties can appropriately halt meta-argument. We will identify three problems for the view. The first is whether it is clear to the disputants in the cases in question who has lost the meta-argumentative



² See Aikin and Casey (Forthcoming b) for an account of the meta- and para-argumentative role of humor in making and commenting on arguments. Justice Elena Kagan (of the US Supreme Court) used this particular image of stuffing rabbits into a hat in her majority decision in *Borden* (2021).

³ See Phyllis Rooney (2010) and Catherine Hundleby (2013) for the claim that argument already too adversarial, even *without* ancillary adversariality.

⁴ Tempest Henning is an exception, see her (2018,2021).

exchange. Being right is not sufficient in argument, as one must make one's point manifest for it to address the controversy. Innocenti's case, by our lights, does not live up to this requirement. The second question is whether the deterrent and affirmative effects of para-argumentative sanction for inappropriate meta-argument yield not only avoidance of further meta-argument, but argument altogether. The adversarial escalation occasioned by name-calling and shouting reverberates beyond the particular cases at issue may stand in the way of re-railing the argument, even if it was supposed to be a corrective to the meta-argumentative derailing. Third, and finally, exceptional cases have a vagueness problem with boundaries. Allowing one-self and others exceptions to generally binding rules has a snowballing problem, and for those on the receiving end of these sanctions, it is unclear what their proper bounds are. We will, in what follows, provide a reconstruction of Innocenti's argument and then pose our challenges.

2 Meta-argument and meta-argumentative misfire

The exceptional cases of meta-discussions Innocenti targets for analysis are what she calls "not-all qualifiers," or NAQs. They are instances of men objecting that not all men are rapists, in the midst of discussions of the fact of widespread sexual violence suffered by women (and overwhelmingly perpetrated by men). Or they are instances of insisting that not all white people are racists, in the process of discussing the consequences of systemic racism against people of color (and overwhelmingly perpetuated by whites).5 These moves are, on Innocenti's analysis, both mistaken and self-serving. They are mistaken because the speaker has confused generics with universally-quantified propositions. So, objecting that there are outliers to these claims is akin to objecting that unplugged refrigerators do not keep food cold, when someone says that refrigerators keep food cold. Or to counter that *lion cubs* do not roar, when posed with the claim that lions roar. The main logical error of NAQs is that they are rooted primarily in holding generics to an inappropriate standard of scrutiny as universals. Further, Innocenti holds that such errors are self-serving because the speaker is asking that they themselves be recognized as one of the exceptions. For this reason, using an NAQ "return(s) attention to feelings of individuals at the center of white supremacist and patriarchal systems" (2022, 356).

The challenge in NAQs, as Innocenti frames it, is that they are meta-argumentative moves, purporting to identify a logical error. In particular, NAQs are accusations of "hasty generalizations about men or white people" (2022, 354). The NAQ challenge, then, as a bid to open a meta-argument is to make a point of order by way of a fallacy charge. The trouble with the challenge, however, is that it is based on a misreading of a generic claim as a universal. But with this misunderstanding and correction, those making NAQ objections make further commitments and enact other roles. Innocenti identifies three.

⁵ See Monique Judge (2018) for a discussion.



First, NAQs are instances of meta-argument "derailing" and "interrupting" a first order argument of consequence (2022, 351). The pause for the meta-argumentative objection and request for clarification is costly, and so is a disruption of the discussion at hand—namely, sexist and racist practices. Second, NAQ objections implicate something about their targets—that they have generalized *hastily*. This, Innocenti counters, is a "ridiculous insult," because activists and those making cases about systemic racism and rampant sexism "are not stupid" (2022, 354). The background assumption is that fallacy charges have implicit accusations of incompetence. Third, and finally, Innocenti argues that those who pose NAQs implicate that *they themselves* are the exceptions that must not only be recognized but positively honored. An NAQ is a "plea for personal reassurance" that the speaker is one of the good ones (2022, 354).

What is important is that some NAQs are instances of illegitimate retreats to meta-dialogues. Thus we propose a friendly amendment to Innocenti's case. As we see it, the NAQ strategy described by Innocenti is an example of the more general class of meta-argumentative fallacies. Let's explain these briefly. As we've mentioned, there are errors of argument in the first order, as when we hastily generalize about lions, or affirm the consequent about climate change. By contrast, fallacies of meta-argument occur when we argue badly about arguments. These fallacies are peculiar, in part for the reasons Innocenti describes: one needs to know, or claim to know, something about arguments to make them. So, in the NAQ case, one presents themselves as knowing how generalizations work, and so volunteers a correction. But the meta-argumentative correction is mistaken and self-serving.

Though our analysis in this paper will focus on Innocenti's account of NAQs, we'd like to note that this kind of meta-argumentative misfire can happen in several different ways. Here we will briefly mention two. In the course of a discussion about the deaths of unarmed Black people at the hands of law enforcement, one might ask what about Black-on-Black crime? Here the meta-argumentative move, called whataboutism, does not raise questions about the nature or frequency of death at the hands of law enforcement, but rather suggests the real problem lies elsewhere, and so the dialogue has gotten off on the wrong foot (Aikin and Casey 2022b). Like NAQs, the whataboutist move is also self-serving, as it rests on recasting a discussion in terms favorable to the objector. It is meta-argumentative, because it consists in characterizing the overall stakes of a dialogue instead of claims within the dialogue. Another case of meta-argumentative maneuvering consists in complaining that all the relevant stakeholders of a debate are not represented, and, if they were, the results of the discussion would tend to a compromise between these views. Call this bothsiderism, for the underlying concept is that the truth, or the most plausible position, will consist in a compromise or overlap of existing views (Aikin and Casey 2022a). Like NAQs and whataboutism, this is self-serving because the objector insists that their view is plausible merely because it is a possible position. It is metaargumentative, because the objector hopes to salvage their view not on the strength

⁶ See Herman and Oswald (2022) for a case that argumentative criticism (in their case, the straw-manning of one's interlocutor) comes to be a criticism also of their intellectual character.



of its claims, but merely because all debates ought to eventually accommodate some aspects of all of the participants or at least address the outliers.

Argument is self-regulating activity—we are all the referees—and so participants' occasional recourse to meta-argument is necessary. But this is what makes meta-argumentative fallacies such as NAQs, whataboutism, or bothsiderism particularly galling; they take the norms of a well-run exchange and turn them inside out, turning what might seem like progress in one sense (the expansion and clarification of argument norms) into a yet another opportunity for obstruction, frustration, and derailment. For this reason, they are especially severe transgressions. How, then, does one deal with them?

Innocenti's view is that "the mechanism of paying a social cost" can halt these metadiscussions. And so she returns to Krabbe's thesis that those who lose meta-argumentative exchanges must pay a price:

Consequently, Krabbe's ... proposal to penalize an interactant who has not "won the metadialogue" as a way of halting retreats to metadialogues deserves to be taken seriously. (2022, 359)

The costs to be paid that Innocenti reviews, again, are the costs of being the target of "shouting, cussing, sarcasm, name-calling." Innocenti favorably surveys a number of confrontational statements from a variety of writers. These are representative:

- Have a seat and pour yourself a nice, big cup of SHUT THE FUCK UP (350).
- May we never have to explain this shit to your sensitive thin-skinned, snowflake asses again (350).
- Well, read a fucking article. Being aware of what is happening in our world isn't difficult (351).
- This is a douche move, and YOU ARE LIKE THESE MEN (356).
- [D]o you want me to give you a cookie for being a 'good man'? Pause for a second to think about how MESSED UP that is (356).

The take-away is that NAQs and other misbegotten meta-arguments are, as Innocenti puts it, "flawed logically ... and relationally" (2022, 359), and so they should be halted and those who propose them must pay the para-argumentative penalty for failed meta-discussion. "[N]ame-calling, cussing, sarcasm, and shouting are designed to display how outrageous using NAQs is and the resentment it deserves" (2022, 359). This expression, then, halts the move to meta-argument.

3 Para-argument and the manifestness problem

Losing a meta-argument requires that one pay a price, and since the only argumentative (or meta-argumentative) price to be paid is as being adjudged as incorrect on the matter (which is equivalent to the loss), the price paid must be in *para-argument*.



Para-argumentative moves are those that can be dependent upon argumentative performances, but themselves are non-argumentative. So, shouting, humor, sarcasm, and enthusiasm are para-argumentative contributions, and one can be penalized or rewarded with these. These are all best considered as para-argument bearing on facial status. Applause after an academic talk, for example, or unhappy muttering during it, or the tone of address are all para-argumentative moves that can be expressions of approval or disapproval. By our lights, these are facially-related, as they bear on an interlocutor's self-image, relational social standing, and their feelings about themselves. The assumption is that para-argumentative penalties paid for meta-argumentative errors will be ones that, though breaking the rules of positive politeness, are in the service of protecting other norms. So, denying acceptance to a discussion, rejecting one as part of a group, having one's preferences explicitly denied or downgraded, are heavy facial challenges that can stand as ways of maintaining the norms of the group. More specifically, the penalties surveyed are explicit politeness-negating strategies, and they deviate from background norms of exchange. These, again, can range from the relatively mild to the very severe: thus, a humorous analogy to rebut an argument (e.g., "the ontological argument is like making a god of words") to the outright insult (e.g., "well, that's a stupid idea, you nitwit"). And again, tone can do much of this work, since one can implicate one's opinion of the interlocutor and their performance by how one delivers a comment like, "Looks like Donny has a point of order." The important assumption here is that the penalty is a penalty only if norms of face-politeness are the defaults. If norms of face-politeness are suspended, then penalties are not penalties, but merely the price of exchanging with others in this argument.

Further, it's worth noting that the para-argumentative moves here have meta-argumentative implications running conversationally upstream. So, in the case of ending a facially-damaging reply to a failed argument challenge with "So, shut the fuck up," the implication is that visiting the para-argumentative sanction on the challenger is evidence that the meta-argumentative bid is foregone and unwinnable. The para-argumentative sanction, if it is to halt the move to meta-argument, must itself function as a stand-in for a meta-argument about the meta-argumentative bid—namely, that the challenger's move to meta-argument will fail.

Our argument here will be that Innocenti's case for escalations in para-argumentative adversariality to halt meta-arguments has three problems. To be clear, we are not defending those who use NAQs or commit other meta-argumentative fallacies, and we agree with Innocenti about all of the errors made and harms enacted in posing them. We also agree that there should be para-argumentative sanction for unfounded and offensive meta-argumentative proposals. The problem, as we see it, is how all of this is communicated. First, the direct use of para-argumentative attacks on face, as an expression of outrage, is not a demonstration that the meta-argumentative proposal is unfounded. This point, we believe is obvious, but it is worth saying explicitly: yelling at someone does not show them that they are wrong. Even when they are. And this point seems more clear when the issue is a meta-argumentative

 $^{^{7}}$ See Aikin and Talisse (2008) for the implicated epistemic background to tone of voice in disagreements.



matter. That is, the person receiving the penalty must also understand that they have not won the meta-argumentative dialogue in order to see the penalty as legitimate. So, however richly deserved, if the penalty is not the result of a demonstration that the meta-argument failed, then it cannot be understood by those penalized as a penalty proper. Without clarity that they have not won the meta-argument, the penalty by the offender's lights is simply abuse. In argument, being right is not enough—one must show how that is the case to those with critical questions. It is with this requirement that we think Innocenti's case fails. In fact, because its objective is to reject and halt the meta-argument, it is an inherent feature of the proposal that it will always fail this objective. The same analysis applies to the other suggestions, in particular those that employ humor or sarcasm. Though humor is a powerful tool at redressing social imbalances and highlighting logical failures among those that are in on the joke, it carries with it the ironic requirement that to find it funny, you have to get it. And the people least likely to get the joke are the one's who are its target (Cochrane 2017).⁸

Let us call this the *manifestness problem*. In order for one arguer to legitimately penalize another for a failed meta-argument, it must be manifest that the meta-argument is a failure. The problem, for Innocenti's purposes, is that NAQs are time-wasting diversions from more pressing issues and occasions for centering non-marginalized people in the midst of discussions of challenges for the marginalized. And so, taking the time to address NAQs gives those who pose them what they are asking for, which seems deeply objectionable. With this, we agree completely. But that is not the end of the story, since every decision under non-ideal or sub-optimal conditions is contrastive—we must also account for the bad consequences of the decision to halt the meta-argument with facial sanction. Only once we've seen the full account of the costs on either side of the decision can it be properly theorized. The problem we see is that in refusing to do the meta-argumentative task of correction, one fails a shared rule of critical discussion, one that runs roughly as a responsibility to defend announced viewpoints from criticisms. Visiting para-argumentative abuse on those with objections (or even questions in search of clarification) without answering the objections not only breaks the rule of defense, but leaves it unclear what the error exactly is to those who have posed the challenge. As David Godden has argued, the obligation to defend rule is not only a norm for managing argumentative escalation, it is also a check on capricious treatment (or perceived caprices) of interlocutors (2021, 854). Ocnsequently, suspending the rule of defense will clearly be received as abusive escalation and a significant break with the norms of well-run argument.

¹⁰ For articulations of the rule of defense as a requirement of reasoned dialogue, and defenses of it against objectors, see: Johnson (2000, 2003), van Eemeren and Grootendorst (2004), Rescorla (2009), Aikin (2021), and Alsip Vollbrecht (n.d.). Taylor argues, further, that the work of de-colonizing philosophy and broader intellectual culture depends on "deepened resources for self-critique" (2015), so requires all living up to the rule of defense.



⁸ For a discussion of the virtuous employment of sneers see O'Brien (2022) and for sneering satire, see Anderson (2022).

⁹ The norm behind this problem follows Johnson's basic rule of argument that it must be *manifest rationality*, so must live up to the rule of defense (2000, 2003).

Much of our response, so far, is posited on the assumption that those posing NAQs have done so sincerely. That is, we've assumed that posing an NAQ is a consequence of being genuinely puzzled about the generic claims, seeing that there are exceptions to them, and posing that challenge. Of course, there are those whose attitudes are not representative of all NAQ challenges, as it is clear that they can be posed eristically and for the sake of making the argumentation more costly in patience and time. As van Eemeren and Houtlosser observe, strategically posed meta-discussions "allow participants to hold up dialogue infinitely by seizing on any opportunity to initiate meta-dialogue" (2009, 205). For sure, given that they are purely strategic and for the sake of undercutting the purposes of the first order dialogue, these are appropriate targets for para-argumentative sanction. But in this case, it is for their insincerity and inappropriate proposal for meta-dialogue, not the content of their claims. 11 So, those who pose NAQs and other meta-argumentative questions may do so sincerely or insincerely. 12 If they do so insincerely, they deserve sanction for their insincerity. And if they do so sincerely, then meeting them with sanction is not yet clearly the answer, since the question is whether the move itself was a winning meta-argumentative bid or not. Those sincerely posing NAQs are the ones under consideration at this stage, since responding with para-argumentative sanction is itself not a demonstration that they have failed at their meta-argumentative duties. Again, assuming sincerity, it is clear that they believe they are posing appropriate challenges.

We should pause to note that the manifestness problem is one that Innocenti's announced normative pragmatic approach needs to address. Elsewhere (2011), Innocenti argues that responses to questionable argumentative tactics should make the relevant norms determinant, make clear how the tactic identified breaks the rule, and show that the speaker is exercising "forbearance" in responding argumentatively instead of jumping directly to sanction or ending the exchange altogether. These tactics she calls "crying foul," and in these, one makes others accountable by articulating what has gone wrong.

Crying foul pressures opponents to repair or abandon questionable tactics by making norms determinate, and by making manifest the badness of opponents' tactics and that the speaker is exercising forbearance (2011, 186).

The key is that one tips the argumentative balance of the exchange by making the badness of the opponent's contribution manifest (2011, 187). And in making this error manifest, we have a path for the opponent to retract their move, revise it, or retreat from their objection altogether. In showing the error explicitly, we plot a path to re-rail a de-railed exchange, and one does so by overtly exercising forbearance.

¹² Walton (2011: 379) distinguishes between paralogisms (blunders) and sophisms (purposeful attempts at deception). The meta-argumentative vice of NAQs can be taken to be both in that the vicious use hides behind either the accusation of a blunder and (or) the, for example, naïve understanding of generics. Godden (2022:46) argues that only the vicious use merits sanction.



¹¹ One mark of the insincerity may be that the use of the NAQ (or other meta-argumentative bid) would be that it persistently is offered, even when answers have been given explicitly in other contexts. In this case, again, the issue is not with the bid but with the insincerity (and disruptive purpose) with which it is posed.

In this regard, we see Innocenti's proposed justification for para-argumentative sanction for NAQs to be less a solution to the manifestness problem than an entrenchment of the problem. Again, argumentative sanction is *sanction proper* only if it is recognized as a response to a rule not followed or a norm broken, but if we go directly to sanction without demonstrating the rule being broken, it is received as abuse. The dilemma for the manifestness problem with NAQs is that one either accedes to the costly and insulting meta-argument and thereby corrects it (and then may sanction), or one prevents the meta-argument with halting sanction but thereby breaks a rule of defense. Innocenti has shown that there are good reasons for the latter, but she has not explained why this option outweighs the costs to argumentative principle and the clarity of what is being preserved. And we have made the case that this cost is (*prima facie*) prohibitively high.

Our second question is whether Innocenti is right that these para-argumentative sanctions have the deterrent and salutary effects of repairing the argumentative exchange she proposes. Innocenti holds that:

[A]ddressees are free to accept social costs they may pay or not pay ... They are not bullying any individual or damaging any particular interpersonal interaction or relationship, but counting on the conspicuous, vivid display of the merits of their position to cultivate a context where addressees can hold each other accountable for recognizing the merits of their demands (2022, 361).

There is a sense that Innocenti is right here, as it is clear that the structure does have those posing NAQs being held to account, but halting the metadiscussion with insults is precisely the opposite of cultivating a context of argumentative accountability. This is because halting the metadiscussion and visiting para-argumentative sanction on those posing a critical question without given justification does not sound like any such community of reply is being cultivated. Innocenti's proposal is that this is done for the sake of better run argumentative exchange, but it seems rather to have made those exchanges unwelcome not only to meta-arguments of this and other forms, but argument generally. For sure, anyone sanctioned for their meta-argumentative bids without being given justification will have grounds to hold that this is, in fact, not a context for evaluating the merits of claims.

We repeat that the argument here has been to concede and agree to all of the upstream evaluations of NAQs and those posing them. Innocenti's analysis of the errors and the intellectual vices that yield them are correct. The NAQ challenges are based on misinterpretations, they unreasonably implicate argumentative incompetence, and they are inappropriate requests for compliments. And given these, NAQs are reassertions of social dominance by those who are ignorant of the discriminatory and sexist systems from which they benefit (2022, 350). And consequently, they deserve censure. Again, with all this, we agree. It is with the rider that one may censure those who make NAQ challenges without showing that they are in error that we disagree.

The point can be put otherwise: we cannot have someone pay a penalty for a failed meta-argument if there is no meta-argument allowed. And the claim that this policy is one that "is entirely in the spirit of a critical discussion" and "cultivate(s) a



context where addressees can hold each other accountable for recognizing the merits of their demands" (2022, 361) seems deeply implausible if not clearly false. Para-argumentative escalation, with raised voice, profanity and obscenity, or ALLCAPS reply show outrage, but they are also tools of preventing conversation and silencing. If manifestness is an objective of argumentation, the ambiguity of these gestures makes them inappropriate stand-ins for argument. They may function *alongside* arguments, but they cannot (especially under conditions of disagreement or significant misunderstanding) play the role of clarifying or truth-directing contributions.

It is worth emphasizing that the para-argumentative program of stopping metadiscussions with facial attack is widely received by recipients *not* as re-establishing argumentative norms. In fact, it is quite the opposite. And the case in point: being shouted at in reply to what at least seems (to them) a legitimate point of order with an NAQ is not interpreted by the targets of the criticism as well-run dialogue. In fact, many take this kind of response to not only be a failed reply to the metaargumentative bid, but as breaking the argumentative context altogether. We quote the following not as endorsements of their positive views but as evidence that Innocenti's proposal that facial attack and overt hostility improves the exchange is likely inaccurate when it comes to exchanges between progressives and conservatives:

If the illiberal feminists were truly confident in their views, they would welcome disagreement and dissent. . . . It's the illiberal left and the illiberal feminists that fear debate (Powers 2015, 171).

It is the left that uses thug tactics to silence voices rather than celebrating the grand panoply that makes up American politics (Shapiro 2013, 258).

[Liberals] self-consciously hold themselves outside the argument and make snippy personal comments about anyone who is actually talking about something (Coulter 2002, 13).

And that's the goal of the authoritarian Left: to cow everyone into silence, except those who agree with them (Shapiro 2021, 133).

And the chilling effects of what conservatives call "cancel culture" are well-documented (as shown by the 2020 Cato institute survey revealing that 52% of self-identified liberals, 64% of moderates, and 77% of conservatives "self-censor" in political discussion, for fear of personal consequences by those who disagree). The point, again, is not to justify the substantive perspectives here, but to show that visiting sanction on others as a way of halting a meta-argument is not received as reinforcing proper argumentative norms or as evidence that their contributions are badly founded. In fact, it is received as undercutting the norms of argument and avoidance of an unrestricted discussion of the issue. Without the argument

¹⁴ See Aikin and Casey (Forthcoming a) for a review of fallacious inferences drawn from this break-down.



¹³ https://www.cato.org/survey-reports/poll-62-americans-say-they-have-political-views-theyre-afraid-share#

(or meta-argument), sanctions are interpreted as mere abuse and refusal of critical exchange.

A related problem with jumping to censure without presenting evidence for the error is that it is not at all clear what analogous applications of suspending standard norms are now made appropriate. In reactionary conservative lines, the suspension of what they take to be politeness norms then licenses a response in kind. Ben Shapiro's take is exemplary of this reasoning:

We've tried to be polite; they've spit in our faces, then blamed us for debasing the level of our national discourse [...] That must end now [...] There is only one way to stop a bully: to punch back. We've seen who the bullies are and what they've done. Now it's time to fight them (Shapiro 2013, 262–63)

Given that the proposed turn to sanction to prevent meta-argument is also a paraargumentative facial attack, it should be clear that it will be too easily interpreted as bullying and without clear boundaries. Are, now, any points of dialectical order cause for sanction without explicit justification? It, unless the reasons are given in the first place, is a plausible interpretation of the reaction. This is because any fallacy charge will interrupt first-order argument, it will impute a common and easily correctable error (and thereby, a level of incompetence), it will be an assertion of intellectual dominance, and it will center the speaker as an arbiter of dialogical norms. So, noting that the other side has attacked a straw man of one's view, or holding that some explanations seem ad hoc, or holding that there are false equivalences all seem to check these boxes. Only if the speaker knew that the NAQ was a special kind of error demanding a special kind of response could this difference be detected. But if the difference were detected by the speaker, then it is considerably less likely that the error would be made in the first place. So, we have, given the assumption that NAQ charges arise because of systematic ignorance, not only reason to think that the sanction visited will not be interpreted as correction, but it will be unclear to interlocutors and onlookers what the relevant norms of correction are.

These unhappy results should not be a surprise, since the explanation for why this group makes meta-argumentative bids with NAQs to begin with is that they are not astute interpreters of the argumentative situation. It is, then, reasonable to expect that they will not have the interpretive approach Innocenti holds they should employ to properly interpret the facial attacks in reply. Were they educated on the issue, they would see the sanction as appropriate and redirecting. But if that were the case, they would not have posed the NAQ. But they are, in fact, not educated on the issue, which is why they posed the NAQ. So they are not going to interpret the sanction in the way that re-rails the conversation. The circumstance is a kind of conditional paradox: were they the kind of person who'd interpret the correction properly, they would not be the kind of person who'd need the censure; and since they need the censure, they are not the kind of person to interpret the censure properly. The result, we hold, is that censure in the form Innocenti proposes is either ineffective or counterproductive.

Our argument has been that Innocenti's proposal to halt meta-argument with paraargument has an *audience problem* (captured by what we'd called the *manifestness*



problem and now with this conditional paradox). Those receiving the censure, given the dialectical situation, will not see it as censure proper. So, even if deserved (and we agree that a wrong has been done that indeed deserves censure), the censure will not function as the normative correction Innocenti holds it does. It will not be heard as a corrective to a culture of argument, but rather as its demise.

However, there are audiences for these communications who can interpret them properly—namely, those who are already sympathetic or in agreement with the views criticized by those proposing the NAQs. That is, the "shouting, cussing, sarcasm, [and] name-calling" may be at those who pose the NAQs, but it is all for those who are already convinced that NAQs are out of order. Innocenti holds that "addressees are free to accept social costs they may or not pay" (2022, 361), and argues that the conditions for those costs are clear because those enforcing them are out to have interlocutors recognize that they have broken rules of exchange. Innocenti argues that these confrontations are "conspicuous, vivid display(s) of the merits of the position" that are reinforced norms of argumentative accountability (2022, 361). But, again, the problem is that this clarity is foregone for those receiving the sanction by halting the meta-discussion. Only those who already know can properly understand these contributions as conspicuous or vivid correction. The result is that it is reasonable to think that the target for the censure will not see it as censure proper, but those already convinced of the correctness of the censure will approve it. This is the audience problem, and given this problem, we have a question: whom does this communication benefit?¹⁵ Assuming that the conditional paradox bears on those targeted by the sanction, it seems more likely that one's ingroup is the preferred audience for these communications. Those sanctions re-affirm the norms for those who already know and accept them by making an example of those who (negligently) fall afoul of them.

4 Objections and replies

The meta-argumentative burden of addressing NAQ challenges is fraught. Here we must address objections to our proposal that halting meta-argument with deterrent sanctions undercuts argumentative norms and is counter-productive. First, there is the problem of the unequal distribution of epistemic labor. Second, there is the problem of potential power and social placement asymmetries that impinge on the meta-discussion. And third, there is the problem of tone-policing we seem to be engaging in with our case. We believe we have appropriate answers to these challenges.

The first objection runs that the distribution of epistemic labor is objectionable with our proposal, since it seems that the burden of educating the politically ignorant

¹⁵ In this, we think that the analysis of moral and intellectual grandstanding is useful—one uses moral and intellectual talk as a means to publicly humiliate and intimidate others and to make oneself feel better and appear more virtuous in others' eyes. It is, then, an abuse of moral and intellectual talk. See Tosi and Warmke (2020) for an account with regard to moral talk, and see Alsip Vollbrecht (n.d.) for an account with regard to moral and intellectual talk.



falls to the disadvantaged. Our reply begins with the observation that on the assumption of a domain of epistemic privilege (namely, that oppressed people understand their oppression better than their oppressors do), there is a minimal sense that the labor distribution is a logical corollary of the distribution of the relevant knowledge. With knowledge comes the burden of informing the ignorant, and this can be best taken as a further element of the alienation that comes with oppression. But the fact of the matter is that the distributions of this labor can be so that *others* can take on this role. Assuming many of these conversations are polylogical affairs (with many onlookers and other discussants), members of non-oppressed groups can take on the task of meta-argumentation with the NAQ contributors. In fact, we might even say that this is precisely the kind of role that white allies and male feminists can play in these discussions—to have sidebars with those who make these inappropriate contributions, including but not limited to NAQ objections.

A second reply to the distribution of epistemic labor problem for our proposal is that the obligations internal to argument are not themselves obligations to argue. Some arguments are hopeless, given the intransigence of the interlocutors or the limits on resources (time, energy, patience). And it may be appropriate in some cases to skip the meta-argument for para-argument, but we should be clear: these are not cases where we are improving argument or the context for the exchange (as Innocenti has proposed), but a case of overtly opting out of it. The oppressed are under no obligation to argue in these cases, and they should take their best means of addressing those who refuse to see the problem. And there may be salutary consequences for taking these routes, but they are not meta-argumentative (or argumentative) results.

The second challenge to our proposal is that meta-argumentative discussions are sites for stark asymmetries that portend epistemic injustices. The problem is that bids for meta-argument open contexts where members of oppressed groups can be given lower credibility than their oppressors. With an NAQ, as Innocenti rightly observes, a member of a privileged group is asserting their privilege by making a point of order and holding up the discussion until it is answered. And with this, the meta-argument can recapitulate the problems that the first-order discussion was out to address—that biases infect our lives, and so they will infect our discussions about those lives, and discussions about those discussions. Burdens of proof, for example, are drawn by asymmetries of power and social placement outside the conversation, and so it stands to reason that those asymmetries will reprise also in the meta-discussion (as noted by Linker 2014: 74, 88).

Our reply is that NAQs are common errors, ones that those who are familiar with the cognitive-argumentative terrain will have seen before. In fact, one reason why it seems that there is a strong inclination to skip meta-argument right to sanction is the impression that it is a discussion one has all too regularly. But if that is the case, there are effective means to address the burden of proof asymmetry—one can point to how posing the NAQ itself is an example of the kind of problem being addressed. If an interlocutor is noting that *not all men*, they are acknowledging the wrongs done by the other men, so there is a sense that the audience for this reply will find these reasons accessible. And note that once this has been identified, the following points should be makeable. Fragile masculinity is a problem that, first, makes material



circumstances for women unsafe, and second, makes discussion of those material situations badly run. White fragility, too, has the same looping effect in that not only are the material conditions for people of color too often unsafe because of racist white anxieties, but bringing those facts up is also unsafe because of white anxieties about their responsibility for these material conditions. If this is an argumentative context, by hypothesis, then showing how the structural problems are recapitulated in the exchange is useful. In fact, that these lines of reasoning are widely rehearsed and clear to those who see them gives us reason to think that that such a reply is stateable with authority. Every teacher, by analogy, knows the four or five points in a lesson that inattentive students regularly misunderstand, and a little experience and practice allows them not only clear but efficient replies to frequently erroneous inferences and challenges.

Further, given that there is already literature on the matter readily available, another option is for a discussant to share with the NAQ objector that their challenge is out of order, given where the broader settled discussion is at this point. They may send the NAQ objector to read one of the articles cited by Innocenti, or to classics in the field, depending on what level of engagement we can expect of our interlocutor. In this regard, we follow Diego Castro's (2022) suggestion that the super-norm for sub-optimal argumentative settings is that of restoring symmetry or compensating for asymmetries. In this case, taking the tone that one is correcting a common error and redirecting the erring party to some educational opportunities is the appropriate compensating para-argumentative move. ¹⁶

The third objection is that our proposal is inappropriate tone-policing of marginalized voices; in particular those expressing outrage over the racism and/or sexism the discussion was supposed to be addressing. Expressions of anger can be wellfounded, since anger has an ethical component (one is expressing a reaction to what one sees as a moral bad), and there are certainly appropriate targets for this reaction. Further, rules that require that the oppressed regulate their anger have the possibility that those rules prompt them to ignore their anger and not take steps to remedy their oppression (as noted by Cherry (2022, 8).¹⁷ Thus, our proposal can be interpreted as a form of *affective injustice*.

Our reply is not that any of the expressions of anger or facial penalties visited are objectionable. We are not saying that arguers in these circumstances must take on a kind of burdened virtue of tolerance and reply with equanimity to meta-argumentative bids with NAQs (on analogy with Tessman's (2005) observation of political virtues that impede individual flourishing, but are useful for the context). Rather, we've agreed that the para-argumentative sanction is appropriate. But it will have expressive and reparative function only if it is visited *after* a winning meta-argumentative exchange. It does not have those salutary effects if it comes before, instead of, or as a way of preventing meta-argument. Arguments from outrage are incomplete arguments. They signal that the speaker judges a wrong has been done, but if they are performed only in the place of making the reasons for the outrage explicit, they offer

¹⁷ See Njovane (2015) for an account of how norms of politeness "reinforce black oppression" (126) if they are designed to curb critique.



¹⁶ See also Aikin and Talisse (2008) and Henning (2018) for epistemic implications of tone of voice.

no reasons beyond that announcement. They are inadequate as arguments, and if they are posed in the stead of arguments (and to prevent further argument) they are destructive of well-run argumentative exchange.

Expressing anger under these conditions of explicitness is useful, since it not only can emotionally fit the damage, but it can be a way of expressing one's dignity in the face of mistreatment (as noted by Cherry 2022, 7). And with this, we agree. But the issue is whether this anger's presentation, without the explicitness/ manifestness condition, is counter-productive in affirming argumentative norms and in re-railing derailed conversation. Innocenti's proposal is that they are effective, and we've argued here that they are not. Further, there is reason to believe they are in fact counter-productive. Crucially, the matter is not whether or how anger to be expressed, but, procedurally, when. Our case is that the matter is not about the anger, but about it skipping ahead of the procedural line. Such is the nature of anger to jump ahead of its appropriate time. But if expressions of anger are tools, the key is to prevent the retributive impulses behind anger expressions from undercutting the reparative and expressive effects one aims for in expressing that anger. Our case has been that maintaining an argumentative context requires that one maintain argumentative norms (particularly, that of defense). So, anger and sanction may be perfectly appropriate for the circumstances, but they cannot be legitimately used to prevent argumentative exchange. They are appropriate only as a result of the exchange.

Declarations

Conflict of interest The authors have no conflicts of interest (financial or non-financial) bearing on this academic work. It is original work, and it is not (and will not be) submitted elsewhere while under consideration at Argumentation.

References

Aikin, Scott. 2020. Dialectical Responsibility and Regress Skepticism. In *Epistemic Duties: New Arguments, New Angles*, ed. Kevin McCain and Scott Stapleford, 268–280. New York: Routledge, 2020. | Series: Routledge studies in epistemology: Routledge. https://doi.org/10.4324/9780429030215-21.

Aikin, Scott F., and John Casey. 2022a. Bothsiderism. *Argumentation*. https://doi.org/10.1007/s10503-021-09563-1.

Aikin, Scott F., and John Casey. 2022b. Straw Man Arguments: A Study in Fallacy Theory. London: Bloomsbury.

Aikin, Scott F., and John Casey. Forthcoming a. "Free Speech Fallacies as Meta-Argumentative Errors." Argumentation.

Aikin, Scott F., and John Casey. Forthcoming b. "Knock Knock: Meta-Argumentative Humor, Who?" Inquiry: Critical Thinking Across the Disciplines.

Aikin, Scott F., and Robert B. Talisse. 2008. Modus Tonens. *Argumentation* 22 (4): 521–529. https://doi.org/10.1007/s10503-008-9103-1.

Alsip Vollbrecht, Lucy Alsip. n.d. "What's Wrong with Intellectual Grandstanding?"

Anderson, Luvell. 2022. Sneering Satire. *Aristotelian Society Supplementary* 96 (1): 269–288. https://doi.org/10.1093/arisup/akac010.

Bermejo-Luque, Lilian. 2010. Second Order Intersubjectivity: The Dialectical Dimension of Argumentation. *Argumentation* 24 (1): 85–105. https://doi.org/10.1007/s10503-009-9130-6.



Borden vs. United States. 2021. Supreme Court of the United States.

Castro, Diego. 2022. Argumentation in Suboptimal Settings. Argumentation. https://doi.org/10.1007/s10503-022-09571-9.

Cherry, Myisha. 2022. Political Anger. Philosophy Compass. https://doi.org/10.1111/phc3.12811.

Cochrane, Tom. 2017. No Hugging, No Learning: The Limitations of Humour. *The British Journal of Aesthetics* 57 (1): 51–66. https://doi.org/10.1093/aesthj/ayw073.

Coulter, Ann H. 2002. Slander: Liberal Lies about the American Right. 1st. Pbk. New York: Three Rivers Press

Godden, David. 2021. The Compliment of Rational Opposition: Disagreement, Adversariality, and Disputation. *Topoi* 40 (5): 845–858. https://doi.org/10.1007/s11245-021-09768-9.

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

Govier, Trudy. 2000. The Philosophy of Argument. Virginia: Vale Press.

Govier, Trudy. 2021. Reflections on Minimal Adversariality. Informal Logic 41 (4): 523–537. https://doi.org/10.22329/il.v41i4.6876.

Hamblin, Charles L. 1970. Fallacies. London: Methuen.

Hample, Dale, Fabio Paglieri, and Ling Na. 2012. The Costs and Benefits of Arguing: Predicting the Decision Whether to Engage or Not. In *Topical Themes in Argumentation Theory*, ed. Frans H. van Eemeren and Bart Garssen, 307–322. Dordrecht: Springer Netherlands. https://doi.org/10.1007/978-94-007-4041-9 20.

Henning, Tempest. 2018. Bringing Wreck. Symposion 5 (2): 197–211. https://doi.org/10.5840/symposion20185216.

Henning, Tempest. 2021. 'Don't Let Your Mouth': On Argumentative Smothering Within Academia. *Topoi* 40 (5): 913–924. https://doi.org/10.1007/s11245-021-09751-4.

Herman, Thierry, and Steve Oswald. 2022. Chapter 5. "You want me to be wrong": Expert ethos, (de-legitimation, and ethotic straw men as discursive resources for conspiracy theories. In *Conspiracy Theory Discourses*, ed. Massimiliano Demata, Virginia Zorzi, and Angela Zottola, 99–120. Amsterdam: John Benjamins Publishing Company. https://doi.org/10.1075/dapsac.98.05her.

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

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

Innocenti, Beth. 2022. Demanding a Halt to Metadiscussions. *Argumentation* 36 (3): 345–364. https://doi.org/10.1007/s10503-022-09569-3.

Johnson, R. 2000. Manifest Rationality. Mawah, NJ: Lawrence Erlbaum.

Johnson, Ralph. 2003. "The Dialectical Tier Revisited." In Anyone Who Has a View: Theoretical Contributions to the Study of Argumentation, 8:41–52. Argumentation Library. Dordrecht: Springer Netherlands. https://doi.org/10.1007/978-94-007-1078-8.

Judge, Monique. 2018. "Not-All-White-People-a-Definitive-Disclaimer." Blog. *The Root* (blog). May 5, 2018. https://www.theroot.com/not-all-white-people-a-definitive-disclaimer-1825835641.

Krabbe, Erik C. W. 2003. Metadialogues. In Argumentation Library, ed. Frans H. Van, J. Anthony. Eemeren, Charles A. Blair, A. Francisca. Willard, and Snoeck Henkemans, 83–90. Dordrecht: Springer. https://doi.org/10.1007/978-94-007-1078-8_7.

Linker, Maureen. 2014. Epistemic Privilege and Expertise in the Context of Meta-Debate. *Argumentation* 28 (1): 67–84. https://doi.org/10.1007/s10503-013-9299-6.

Njovane, Thando. 2015. "The Violence Beneath the Veil of Politeness: Reflections on Race and Power in the Academy." In *Being at Home: Race, Institutional Culture and Transformation at South African Higher Education Institutions*, edited by Pedro Alexis Tabensky and Sally Matthews, 116–29. Pietermaritzburg, South Africa: University of KwaZulu-Natal Press.

O'Brien, Lucy. 2022. 'Sneering, or Other Social Pelting.' *Aristotelian Society Supplementary* 96 (1): 245–268. https://doi.org/10.1093/arisup/akac011.

Paglieri, Fabio. 2009. "Ruinous Arguments: Escalation of Disagreement and the Dangers of Arguing." In Argument Cultures.

Paglieri, Fabio, and Cristiano Castelfranchi. 2010. Why Argue? Towards a Cost-Benefit Analysis of Argumentation. Argument & Computation 1 (1): 71–91. https://doi.org/10.1080/194621609034945 84.

Powers, Kirsten. 2015. The Silencing: How the Left Is Killing Free Speech. Washington, D.C: Regnery Publishing.



Rescorla, Michael. 2009. Epistemic and Dialectical Regress. *Australasian Journal of Philosophy* 87 (1): 43–60. https://doi.org/10.1080/00048400802215653.

- Rooney, Phyllis. 2010. Philosophy, Adversarial Argumentation, and Embattled Reason. *Informal Logic*. https://doi.org/10.22329/il.v30i3.3032.
- Shapiro, Ben. 2013. Bullies: How the Left's Culture of Fear and Intimidation Silences America. New York: Threshold Editions.
- Shapiro, Ben. 2021. The Authoritarian Moment How the Left Weaponized America's Institutions against Dissent.
- Taylor, Paul. 2015. "Towards a Decolonial Analytic Philosophy: Institutional Corruption and Epistemic Culture." In Being at Home: Race, Institutional Culture and Transformation at South African Higher Education Institutions, edited by Pedro Alexis Tabensky and Sally Matthews, 203–20. Pietermaritzburg, South Africa: University of KwaZulu-Natal Press.
- Tessman, Lisa. 2005. Burdened Virtues. Oxford University Press. https://doi.org/10.1093/0195179145. 001.0001.
- Tosi, Justin, and Brandon Warmke. 2020. *Grandstanding: The Use and Abuse of Moral Talk*, 1st ed. Oxford University Press. https://doi.org/10.1093/oso/9780190900151.001.0001.
- van Eemeren, Frans H., and Peter Houtlosser. 2004. A Systematic Theory of Argumentation. Cambridge: Cambridge University Press.
- van Eemeren, Frans H., and Peter Houtlosser. 2009. How Should One Respond to Fallacious Moves? Argumentation and Advocacy 45 (4): 198–206. https://doi.org/10.1080/00028533.2009.11821708.
- Walton, Douglas. 2011. Defeasible Reasoning and Informal Fallacies. Synthese 179 (3): 377–407. https://doi.org/10.1007/s11229-009-9657-y.

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

Springer Nature or its licensor (e.g. a society or other partner) holds exclusive rights to this article under a publishing agreement with the author(s) or other rightsholder(s); author self-archiving of the accepted manuscript version of this article is solely governed by the terms of such publishing agreement and applicable law.

