

**“That’s Unhelpful, Harmful and Offensive!” Epistemic and Ethical Concerns with Meta-argument Allegations**

Author

Breakey, H

Published

2020

Journal Title

Argumentation

Version

Accepted Manuscript (AM)

DOI

[10.1007/s10503-020-09538-8](https://doi.org/10.1007/s10503-020-09538-8)

Rights statement

© 2020 Springer Nature Switzerland AG. This is an electronic version of an article published in Argumentation, 2020. Argumentation is available online at: <http://link.springer.com/> with the open URL of your article.

Downloaded from

<http://hdl.handle.net/10072/396705>

Griffith Research Online

<https://research-repository.griffith.edu.au>

# “That’s Unhelpful, Harmful and Offensive!” Epistemic and Ethical Concerns with Meta-Argument Allegations<sup>1</sup>

Hugh Breakey

## Abstract

“Meta-argument allegations” consist of protestations that an interlocutor’s speech is wrongfully offensive or will trigger undesirable social consequences. Such protestations are *meta*-argument in the sense that they do not interrogate the soundness of an opponent’s argumentation, but instead focus on external features of that argument. They are *allegations* because they imply moral wrongdoing. There is a legitimate place for meta-argument allegations, and the moral and epistemic goods that can come from them will be front of mind for those levelling such allegations. But I argue there is a dark side to such allegations, and their epistemic and moral costs must be seriously weighed. Meta-argument allegations have a concerning capacity to derail discussions about important topics, stymieing argumentational interactions and the goods they provide. Such allegations can license efforts to silence, punish and deter—even as they provoke the original speaker to retaliate in kind. Used liberally, such allegations can escalate conflicts, block open-mindedness, and discourage constructive dialogues. In response, I defend “argumentational tolerance”—a principled wariness in employing meta-argument allegations—as a virtue of ethical argument.

---

<sup>1</sup> This document is the postprint version of: Hugh Breakey (Forthcoming) “‘That’s Unhelpful, Harmful and Offensive!’ Epistemic and Ethical Concerns with Meta-Argument Allegations’ *Argumentation*. The final document is available at <https://protect-au.mimecast.com/s/qjlcC1WZxYTMn8IDXhLRRIM?domain=rdcu.be>

## Introduction

In 2016, the Australian government held a non-binding plebiscite on expanding the legal definition of marriage to include homosexual unions. While the opposition endorsed the reform, they objected to the plebiscite on several grounds—including the perceived harms of the ensuing debate. As the opposition leader, Bill Shorten (2016, 415), put it:

Putting the question of marriage equality to a national vote risks providing a platform for prejudice and a megaphone for hate speech... The idea of young people, perhaps yet to come out, seeing the legitimacy of their identity debated on the national stage—that is not an ordeal which we should inflict on any citizen when we have a better path. ...if one child commits suicide over the plebiscite, then that is one too many.

This concern remained prominent throughout the national debate, with continued protestations that the debate itself was a mistake, given its likely harms (Karp 2017). This laid the ground for allegations that conservative arguments against the reform were not just unsound and illogical—but morally wrongful in themselves.

This article explores the epistemic and ethical issues raised by such “meta-argument allegations” (as I will term them). Meta-argument allegations can target the overall conclusion of an argument, a specific line of argument, or a single statement in the course of an argument. Such allegations centre on protestations that an interlocutor’s argument is wrongfully offensive, or that it is likely to trigger unwanted downstream consequences. These manoeuvres are *meta*-argument in the sense that they are not objections interrogating the

soundness of an opponent’s argumentation, but are external commentaries that apply irrespective of such soundness.<sup>2</sup> They are *allegations* because they assert moral wrongdoing.

Calling into question the propriety of meta-argument allegations may seem strange. After all, presuming the allegations are made in good faith, the allogger is simply drawing attention to what is from their perspective a manifestly relevant moral concern. While levelling such an allegation may be confronting, doing so does not automatically breach any serious moral norm (e.g., it is not dishonest) or epistemic standard (e.g., it is not logically fallacious). As such, the practice of levelling meta-argument allegations is performed publicly and even righteously, and—as we will see—is employed across both sides of the political divide.

While accepting that meta-argument allegations can have significant epistemic and moral benefits, I argue that they can also exert a damaging influence on public deliberation and political discourse. I aim to establish the principle of “argumentational tolerance”—understood as a principled wariness to employing meta-argument allegations—as a norm of ethical argument.

My argument begins in Section 1 where I outline the potential epistemic and moral goods offered by meta-argument allegations. Section 2 then explores the two key meta-argument allegations in more detail, explaining their typical effects on argument and the epistemic

---

<sup>2</sup> ‘Meta-arguments’ refer broadly to all arguments about arguments, including the exploration of argumentation methods and standards, such as by argumentation theory (see Finocchiaro 2013, Ch. 3). However, for our purposes here, I use the adjective ‘meta-argument’ to specify a special type of challenge to arguments—namely, one that eschews interrogation of ‘internal’ argumentation qualities like soundness and validity, and instead directs attention to other properties of the target argument, such as its effects in a given context.

concerns these raise. Section 3 discusses how these effects attach to important moral concerns.

As we proceed, I will make use of the following imaginary figures: “Prue”, our protagonist, who makes the initial contribution to the argument, and “Andrew”, Prue’s antagonist, who responds to her contribution. Here and throughout, by “argument”, I refer to any case where Prue and Andrew have a disagreement (or other query) on an issue, and are exploring that issue on the merits, gauging what rationally may be said for or against it and providing reasons to convince the other of the justifiability of their standpoint (Eemeren and Grootendorst 2003, 2). The issue in question may be a moral principle, a policy, a proposed action or an empirical claim. An interlocuter’s purpose may be purely to persuade others of their position’s cogency, or it may be a genuine shared exploration, with each speaker in principle open to persuasion (Walton 1989, 174). Either way, the aim is rational understanding and/or persuasion. Prue and Andrew may come together at a particular time and place—such as a policy debate at a town hall or a structured discussion in a political science tutorial. A shared venue is not essential. Argument may occur online, or (say) across iterated publications in an academic journal. I will assume that the argument is at least somewhat public, and that it touches on issues of social or political significance, which is usually the case when meta-argument allegations are involved.

Before beginning, two scope qualifications. First, the following exploration only concerns allegations of harm or offence when they occur in the specific context of argument—and not in artistic, comedic, bargaining or other contexts.

Second, the problems I will link with meta-argument allegations do not arise simply from their status as *meta*-argument. There are many forms of meta-arguments and meta-argument critiques that are not problematic (Cohen 2001, 79-81, 2009, 60-62). In particular, meta-

argument allegations must be distinguished from meta-argument proposals and meta-argument objections. Meta-argument *proposals* occur when it is possible to raise—in a respectful and compromising spirit—suggestions about mutually mitigating a debate’s potential harmfulness or offence, perhaps by negotiating acceptable ground rules for shared civility and diplomacy (Risse 2000, 25; Ceva 2011, 163). Alternatively, if during an argument one’s opponent says something inappropriate or unhelpful, it is sometimes possible to point this out in a constructive, fallibilistic and non-blaming way. Such meta-argument *objections* are usually directed at smaller wrongs, such as when the interlocuter has just made a mistake (e.g., misgendering someone) or stumbled into an unfortunate manner of expression, from which they can quickly resile. Meta-argument proposals and objections often can be successfully employed, especially when there is an existing context of trust and amiability. They usually avoid disrupting the main argument and can improve standards of civility and mutual respect. In contrast, meta-argument *allegations* are done in an explicit, public and reproving way, usually in direct response to an offending statement, and carrying the implication that the speaker has behaved wrongfully.

## **1. Epistemic and Ethical Goods of Meta-Argument Allegations**

Meta-argument allegations can deliver both epistemic and ethical goods.

In terms of epistemic goods, the most obvious case will occur when our speaker, Prue, had no idea that what she was saying could be harmful or offensive. Even if Prue doubts whether her argument *really is* harmful, or *should be* found offensive, this awareness of others’ concerns can still be valuable information for her—for example, to empower her in future to deliver her argument in a way that avoids alienating or marginalizing others. The meta-argument allegation might also prompt Prue to consider her argument from an alternative perspective, enhancing her understanding of others’ ways of thinking. More deeply, the meta-argument

allegation might draw Prue's attention to the political nature of her argument, and to consider whether irrelevant influences (like her place of birth or schooling), implicit biases or unnoticed assumptions played a worrying role in creating her position (Vavova 2018, 135, 140).

Sometimes the meta-argument allegation may demand that Prue develop stronger evidence and reasons for her view. For example, an allegation of harmfulness may take the form that Prue's argument would be harmful if, a) people believed it, and, b) it turned out to be substantively false or unsound. For example, given that claims about the dangers of vaccination in the face of a deadly outbreak could cause thousands to die—if the dangers were inflated but the claims believed—a meta-argument allegation might demand that it is irresponsible to air such a claim unless there is particularly strong evidence for it. Equally, a meta-argument allegation might assert that Prue's argument is offensive because it demonstrates her flagrant lack of knowledge about the topic, perhaps including a lack of engagement with those who have lived experience on the matter. Both these types of meta-argument allegations might force Prue to better defend her position by developing stronger evidence and better arguments. (Because these two types of meta-argument allegations, at least in part, direct attention back to the substance of the initial argument, they avoid many of the serious 'derailing' effects I explore below. For this reason, I will bracket them from the following analysis, concentrating only on allegations that are entirely 'meta'—applying irrespective of an argument's soundness).

Finally, the meta-argument allegation may have epistemic benefits for the wider community. It can raise awareness of what was said and who said it. This information may be valuable, for example by informing the community's judgment about the alleged offender's character.

While there are some epistemic gains from meta-argument allegations, it is arguably their ethical significance that provides the strongest reasons for levelling them. A meta-argument allegation of harmfulness might help mitigate any ensuing harms, or it may be a critical opening step to begin such harm mitigation (such as by closing the dialogue). This will be especially important if the ethical stakes are direct and high, such as if the speech incites violence.

To be sure, many of these epistemic and ethical goods could be achieved by the more conciliatory meta-argument objections and proposals. However, because it carries the explicit assertion of wrongdoing, the meta-argument allegation provides social censure and highlights to the speaker the perceived moral gravity of her wrongdoing. This effect may be intrinsically valuable as an appropriate punishment for wrongdoing. It may also help restore the social status of those wronged by the offending remark (such as a racist slur). The attaching of the social cost may also have beneficial consequences, such as by contributing to a more constrained—and so more civil or inclusive—future discourse.

Given this array of epistemic and ethical goods, there can be little doubt that in at least some situations, meta-argument allegations are ethically appropriate. In particular, when the risk of harm is direct and grave—such as in an incitement to violence—or the offence given arises from an attack on a subject’s basic humanity and moral standing—such as with the use of a racial epithet—then the magnitude of the ethical concern demands response. When it comes to, for example, racist hate speech, the concerns with harm can become so substantial (Maitra and McGowan 2012, 5-6; Tirrell 2012, 196), that the speech is worthy not only of social condemnation but—arguably at least—legal prohibition (UN General Assembly 1965).

Outside such egregious cases, however, it is an open question whether the achieved goods will outweigh the unwanted costs—a subject to which we now turn.



## **2. Two meta-argument allegations: Epistemic and argumentation concerns**

This section describes the two meta-argument allegations in more detail and explains the typical consequences of their introduction into an argument.

### **2.1 Offence**

Suppose, during an argument, that Andrew alleges that Prue's line of argument, her conclusions, or her language, is offensive. Indeed, Andrew holds that Prue's argument is substantially offensive. She is guilty of wrongdoing and, in order not to be judged blameworthy, she must undertake swift acknowledgement, retraction, apology and efforts at improved performance.

The offence claimed may be to Andrew himself, or to a member of the audience, or to an absent third party. Prue's speech might be alleged to be deliberately inflammatory, but often the allegation will be that—irrespective of Prue's intentions—her speech is offensive.

Sometimes, contextual features might render the speech offensive, such as its timing, if it occurs on a symbolic day, or on the heels of some disaster. Alternatively, the speech may be deemed offensive because of the speaker's membership (or non-membership) of a particular group. For example, a biting critique of group X might be acceptable if it comes from a member of X, but not from an outsider.

The breadth of these possibilities for offence should not be surprising. In any politically, culturally and religiously diverse society, people will cherish different values, embrace different traditions, and identify with different groups. They may also attribute different meanings to morally loaded terms. For example, conservatives and progressives attach different meanings to the term "racism" (Hochschild 2016, 147), meaning that a conservative may misinterpret a progressive's use of the term. These differences would provide myriad

opportunities for offence-giving and offence-taking even if discussants were careful and objective in their appraisals. But, of course, the way we respond to others' words is conditioned by our pre-existing emotional stance towards them—whether we “lean in” towards them sympathetically or not (Haidt 2012, 80). Such emotional responses colour not only our appraisal of their argument's soundness (Gampa et al. 2019, 7), but also our judgments about their argument's offensiveness. As a result, in contemporary political discourse, outraged allegations that the other has said something offensive tumble from both sides of politics. Indeed, in a recent discussion on university censorship, scholars on both sides of the dispute agreed that virtually all utterances offend someone (Cross and Richardson-Self 2019, 7). Yet, rather than recommending generic, neutral rules that might improve the civility of civic discourse, each political group tends to claim offence for specific issues that would muzzle their opponents but leave their own side unchecked.

For example, conservatives may allege that opponents have offended them by paying insufficient respect to patriotic, historical or traditional matters, the mainstream religion, the sacrifice and courage of the country's armed forces, or victims of crime or terrorism. They may take particular umbrage when faced with what they see as progressives' elite condescension and detachment from the common citizen, and they may take offence at the politicization of, for example, severe weather disasters or mass shootings, feeling that it exploits the victims for rhetorical purposes when their political opponents draw connections with climate change or gun control respectively. Such claims of offence can arise in many different contexts, but when they take place within arguments, they usually constitute meta-argument allegations.

For its part, the political left is no stranger to taking offence. Concerns with microaggressions, political correctness and the protection of safe spaces are all bases on which progressives may find offence in conservative arguments. ‘Callout culture’ describes a

social context where offences routinely called out—both generally and in the context of argumentation—and these can lead to further responses like deplatforming or ‘cancelling’.

Equally, the left decries the politicization of criminality and terrorist attacks, and can be offended when conservative arguments draw links from these to immigration or religion.

Nor are these myriad options for taking offence only available across political divisions. They can also occur in specific groups and movements, where offence might arise if, for example, a person’s arguments imply they speak for the group, or if they sideline others’ ideas and experiences (Goldberg 2014).

What are the likely effects when Andrew claims offense at Prue’s words? Sometimes Prue will be amenable to Andrew’s intervention, and this may result in the epistemic and moral benefits noted in Section 1. However, in cases where Prue does not offer swift apology, the usual result is to create an entirely new line of disputation about whether Prue’s speech was indeed offensive. This new debate is a meta-argument: a dialogue about the dialogue. For several reasons, this move tends to derail argument about the original topic.

First, claims about offence assume immediate priority. The argument is no longer about a worrying moral issue. It *is* the worrying moral issue. Because the wrongdoing is happening *now*, within the argument itself, a swift response might be thought to ameliorate the problem—for example, by shutting down the debate, or limiting Prue’s access to an audience. After all, Andrew will not want to “reward” Prue’s offence-giving by engaging constructively with her while she is behaving wrongfully. As well, the evidence of the deed (Prue’s speech) is often visible and a matter of public record. This can make third-party judgment on this issue simpler than on the original topic, which might require subtle evaluation of complex evidence. Furthermore, it makes good procedural sense to settle the rules of argument before engaging in the argument (as Socrates used to do: Plato 1997/380BC, 457c-58d).

Second, this intrinsic priority matters because there is little reason to expect quick agreement on the ensuing disputes. Concerns about offensiveness reflect deep-seated views about the appropriate objects of social vulnerability, respect and sacredness, and the proper extent of tolerance. These views can directly implicate issues of identity politics (on both sides of the political divide (Chua 2018, 166)), about which agreement can be at least as difficult as the original issue under debate.

Third, Andrew's allegation inevitably triggers a further tumult of debates, about which consensus is equally unlikely. Should Prue be made to apologize? Even if she does apologize, should punitive action still be taken? Should her (NGO, government, corporate) institution be pressured to remove her from her role? Should the medium or venue for Prue's speech be pressured to prevent a repeat case? At the limit, should Prue's offensive speech be deemed illegal, and referred to state authorities?

Because these new areas of dispute are far removed from the original topic, they will implicate new areas of knowledge, experience and expertise. This means that if Prue has special knowledge of the original topic, perhaps acquired through long years of academic research, unique personal history, or professional experience, she must now navigate an entirely different discursive terrain before she is able to return to the topic upon which she can make her informed contribution.

In sum, the meta-argument allegation of offensiveness tends to derail discussion about the original topic by introducing new lines of dispute that carry a natural priority, which, at the same time, are unlikely to be easily resolved.

## **2.2. Harmfulness and Unhelpfulness**

The second meta-argument allegation occurs when Andrew charges that Prue's speech is "unhelpful", "counter-productive", "divisive", "harmful" or "damaging". While Prue's

speech might be legitimate from an abstract argumentational point of view, Andrew alleges that it will nevertheless tend to produce some deleterious social phenomenon, such as unhelpful complacency, civic mistrust, hate, bigotry, marginalization, suicide, violence, harmful impacts on the social fabric, damage to the economy or consumer confidence, or weakened national capabilities (such as in a war effort).

There are four ways Andrew might conceive Prue's words as being harmful. First, the envisaged outcome may happen through *sympathetic response*, where the concern is that Prue's speech will help legitimize some unwanted social phenomenon, perhaps by offering succour to those with extreme and antisocial political views and giving them greater confidence to act (perhaps violently or hatefully). Second, the envisaged response may be *antagonistic*. Here, Andrew worries that Prue's inflammatory speech will spark a retaliatory response that might constitute or foment social unrest or political violence. (The two categories of meta-argument allegations are not mutually exclusive. They overlap in cases where Prue's offense-giving can lead to harm, such as retaliatory violence.) Third, Andrew might be concerned about the *institutional* consequences of Prue's speech, feeling that Prue cannot hold her views, and give public voice to them, without that impacting upon her fulfilling her role within a larger organization with its own mandate and branding. A fourth possibility arises if Andrew fears that Prue's words might be taken out of context by a third party and *deployed as evidence* in some harmful way. For example, Andrew and Prue might both belong to social group X, and Andrew fears that Prue's public criticism of X (while perhaps legitimate on its own terms) would nevertheless provide ammunition to X's enemies. In this case, Andrew might appeal to Prue's in-group loyalty, entreating her to eschew her potentially harmful public criticism, to ensure that the besieged group presents a united front. In each case, Andrew's allegation may be that Prue's speech carries a significant risk of harm, or at least that it carries an *unnecessary* risk of harm. The latter allegation might appear

more reasonable, as Andrew is not ruling out all potentially damaging speech, but only speech that needlessly risks harm. (In the Australian example with which we began, the opposition leader stressed alternative pathways to legal reform that would avoid the unnecessary harms of rancorous debate (Shorten 2016, 415).) In actuality, however, it will prove perilously easy for Andrew to judge Prue as being unnecessarily harmful if he is *already* convinced she is wrong substantively. If Andrew thinks Prue is mounting a poor argument for a false conclusion, then he is bound to see any resulting harms as gratuitous. Prue, needless to say, will see things differently.

Just as we saw earlier with offence, Andrew's allegations of harmfulness will derail argument on the original topic by creating an entirely new line of disputation about the social effects of Prue's words, implicating completely different areas of expertise. Once again, Andrew's judgments about harmfulness are bound to be coloured by his pre-existing emotional stance towards Prue and her "side". And, once again, the ensuing disputation will assume a natural priority. If Andrew perceives Prue's words as harmful, then it will make sense to immediately address that issue. Andrew may begin considering what means he possesses to ameliorate that harm—many of which will involve limiting Prue's access to an audience, or at least bringing the debate with her to a close. Not only can the speaker herself be shut down, but pressure can be applied to any person, group, venue or media that reports on or facilitates Prue's current or future speech. These can be criticized as giving a "platform" or "megaphone" to Prue, and thus responsible for amplifying her words and their damage.

As well as assuming priority, the debate over the harmfulness of Prue's words is unlikely to be easily resolved, for it opens a hornet's nest of new disputations. For one thing, the predictions themselves might be controversial. Setting aside cases that implicate very direct and/or immediate harms (such as racist hate speech, noted earlier), predictions about social outcomes are rarely straightforward. Indeed, they can be subject to systemic biases. The

much-researched ‘third person effect’ in communication studies refers to the tendency for individuals who are members of an audience that is exposed to a persuasive communication to expect the communication to have a greater effect on others than on themselves (Davison 1983, 3). This over-estimation of the effect on third parties impacts behaviour, as attitudes to censorship correlate with the message’s perceived influence on others (Chung and Moon 2016, 330). The third person effect could thus drive Andrew to over-estimate the influence Prue’s argument will have on others, and so over-estimate the harms that would follow from that influence.

Unfortunately, there are major complications even in cases where the prediction of harm has been made by an impartial expert backed up by clear evidence. For example, criminologists might be confident about the knock-on effects of certain speech on hate crimes, and psychologists can make evidence-based predictions about effects on emotionally vulnerable people. Yet even in such cases, there remain many factors that make an ensuing meta-argument allegation open for dispute. For example, the (dis)value of the predicted consequences may be differently prioritised by different ethico-political viewpoints. In addition, a given utterance is likely to have multiple potential effects, some of which might be necessary or worthwhile, even as others are problematic. As well, the envisaged harm will usually only occur because other people—other moral agents—make decisions about how they will respond to the speech. If so, then it may be that those moral agents bear responsibility for any resulting harm, and that social pressures would be better aimed in their direction. Finally, the meta-argument allegation may *itself* be thought to have worrying consequences (which we will consider in Sect. 3).

Ultimately, these many distinct factors are all matters upon which reasonable people can disagree—and upon which people with different political perspectives can be expected to disagree. As John Stuart Mill (1859/2003, 21) observed: “The usefulness of an opinion is

itself a matter of opinion: as disputable, as open to discussion, and requiring discussion as much, as the opinion itself.” While the meta-argument allegation of harmfulness can present itself as an external, objective meta-critique of an opponent’s argument, the allegation is bound to be based on claims that are themselves very much open to argument.

In short, there are myriad areas of dispute elicited by a meta-argument allegation of harmfulness. Disputants may disagree not only on the risk that certain speech will create a given effect and that effect’s moral significance, but also on the moral relevance of other agent’s decision-making in creating that effect, and on the countervailing significance of Prue’s entitlements to voice her argument.

### **2.3 Reactance and retaliatory allegations**

Once the meta-argument allegation has been levelled, the ensuing discussions about harm and offensiveness might be explored rationally and constructively. While context will be critical here, there are several reasons why a dispassionate discussion is unlikely. For one thing, the types of argument that Prue needs to rebut the allegations might be open to the self-same allegations. For another, once Andrew alleges Prue is guilty of wrongdoing, and she rejects his allegation, then a likely result is that Andrew’s speech shifts from engaging *with* Prue to speaking with a wider audience *about* Prue. In this shift from an interaction with a peer in equal standing, to an appeal to third parties about what should be done with a wrongdoer, dialogue gives way to monologue, and the argument between the interlocuters collapses (Blair 2012, 158-161; Johannesen 1971, 379-380). For this reason, the meta-argument allegation’s accuracy and impartiality is unlikely to be dispassionately explored.

In addition, meta-argument allegations tend to produce *reactance*. Reactance occurs (inter alia) when one person becomes aware of another’s attempts to control them. The strongest cases concern imposing limitations on people’s freedom, but reactance can also arise through



attempts at influence and persuasion (Coppola and Girandola 2018, 477). Reactance issues in anger, rejection of the influence and even ‘persuasive boomerang’—a perverse shift towards a stronger commitment to the opposite of the aspiring persuader’s message (Kim et al. 2017, 936). Of course, not all attempts at control or persuasion trigger reactance. In ordinary cases, the norms of argument (see Sect. 3.2 below) prevent the use of sheer authority, manipulative rhetoric and external pressures such that any change of mind occurs only through methods agreeable to the person influenced, and only in a mutual manner, by which each has reciprocal opportunities to persuade the other (Keller and Brown 1968, 75). When influence is genuinely consensual and mutually possible, reactance arising from fears of manipulation diminishes.

But meta-argument allegations jolt interlocuters out of this influence-permissive environment. Andrew’s allegation is a direct expression of authority (the authority to determine what is wrong, and to apply social sanctions on that basis) and it attempts to directly foreclose a freedom (‘you cannot say that’) that was not only available to Prue, but being actively employed by her. In such a case, reactance theory predicts that Andrew’s assertion of authority and attempt at control is likely to lead to hostility, vitiating any chance of Prue opening herself up for persuasion. Indeed, reactance is likely to motivate Prue to continue to assert herself against the imposition and—perversely—to value her freedom to do so even more than she did before the meta-argument allegation was raised.

For these several reasons, meta-argument allegations tend to collapse rational argument, auto-immunising themselves from dispassionate, principled exploration and making it hard to expose when they are employed in strategic, biased or mistaken ways.

The prospects for continued argument are even worse when we consider that Prue’s retaliation against charges of offensiveness and harmfulness might take exactly the same

form. Once Andrew has levelled his allegation, Prue may recursively apply the same rhetorical manoeuvre, moving the debate even further away from the original topic. After all, people often take offence at the charge that they have given offence (Kriesberg 2007, 73). Disputants can also think it is harmful and unhelpful to call others' speech harmful and unhelpful. Because Andrew's charges of harmfulness can give rise to efforts at curtailing Prue's speech, she has additional evidence that Andrew's response is socially problematic, as it works to silence (what from Prue's perspective are) reasonable discussions about good faith positions. As a result, the argument threatens to descend into both sides trading accusations that the other is behaving wrongfully simply by participating in the argument. Predictably, this retaliation happened in the example with which we began, with the Australian Prime Minister responding that, in making his allegation, the opposition leader, "demeans the people he claims to represent. He demeans their civility. He insults them. He disrespects them." (Lewis 2016). Needless to say, with both sides levelling meta-argument allegations before the national debate even began, the stage was not set for constructive public deliberation.

## **2.4 A comparison with fallacy and rhetoric**

The use of suspect rhetoric and fallacious reasoning in political discourse is well-known. Yet there is a sense in which the prospects for continued policy deliberation are worse in the case of meta-argument allegations, because of the extent to which they derail discussion by changing the subject.

Consider first some well-known fallacies. Suppose Andrew makes belittling remarks about Prue's character, and Prue takes issue with these. In exploring whether Andrew has committed the fallacy of *ad hominem*, the discussion will focus on the relevance of Prue's character to her case. This discussion may help clarify the overall argument. For example, the ensuing debate might draw out Prue's implicit premise that she is a special authority on the

subject—helping her understand that by appealing to her authority, she has made relevant claims about her status as a knowledgeable and objective expert (Eemeren et al. 2012, 353). Thus, while objecting to an *ad hominem* argument and levelling a meta-argument allegation of offensiveness may seem similar, they are not. The former directs attention back towards what is relevant for the original argument. The latter brackets the original argument and triggers a new discussion.

Similarly, consider if Andrew performed the straw-person fallacy by misrepresenting Prue's arguments. Exploration of whether Andrew has spoken fallaciously would involve the pair developing a clearer understanding of Prue's position (Fogelin and Sinnott-Armstrong 2001, 79). In both these cases, an enquiry into the substance of Andrew's fallacious assertions helps clarify the original argument.

The same is true for much political rhetoric. Consider Albert Hirschman's influential typology of rhetorical arguments used by proponents and opponents of reform. Hirschman (1991, 43, 84) argued that conservatives opposing reforms would be likely to wheel out several stock rhetorical strategies, such as the Jeopardy Thesis (the proposed change involves unacceptable costs) and the Futility Thesis (the proposed change will be ineffective). While these theses may be false or asserted in bad faith, both of them—and, indeed, all of Hirschman's list of conservative and progressive rhetorical ploys—are, at least, *on point*. They all introduce putative facts or concerns that are manifestly relevant to the issue at hand. Responding to those putative facts or concerns will help the discussants and their audience learn more about the issue. For example, discussants will have to consider the evidence that the reform will be ineffective, or that it will have unintended consequences.

A similar point holds true with respect to other common rhetorical ploys, such as the use of metaphor, metonymy, synecdoche and irony. For example, arguing about the appropriateness

of a figure of speech (e.g., calling a welfare initiative a “safety net”) involves a discussion of the phenomenon’s relevant features. The discussion about the analogy’s appropriateness thus remains a discussion about the original issue.

In contrast, as we saw above, meta-argument allegations of offence and harmfulness typically trigger a dispute tangential to the original issue. Combined with their natural priority, such allegations derail argument more profoundly than ordinary fallacies and stock rhetorical manoeuvres.

## **2.5 Epistemic and Cognitive Concerns with meta-argument allegations**

Argument—especially high-quality argument—offers the promise of important epistemic and cognitive goods (Aberdein 2010, 173; Aikin and Clanton 2010, 411-412; Cohen 2009, 52-53). As well as tending to produce increased knowledge and the removal of error, principled argument can deliver increased self-understanding, improved justifiability and coherence of beliefs and values, and strengthened awareness of others’ positions. In some cases, argument can also lead to deeper changes, such as increased fallibilism about one’s views, and the recognition that one’s opponents may have reasonable or legitimate positions.

Meta-argument allegations can threaten the delivery of all these goods. The increased conflict, reactance and defensiveness that meta-argument allegations tend to trigger make disciplined, open and frank argument unlikely. In addition, meta-argument allegations’ derailing effects mean that interlocutors are unlikely to achieve epistemic and cognitive goals in their initial area of exploration, which will typically be an area of need, expertise or interest.

## **3. Moral concerns with meta-argument allegations**

Section 2 explored the predictable argumentational and epistemic consequences of levelling meta-argument allegations. These consequences may give Andrew reason to think that

making the allegation may not prove a straightforward, constructive intervention. But do those consequences give him any *moral reason* to pause?

There are several ways the foregoing considerations attach to areas of moral significance. Some of these (Honesty and strategic use) apply only in certain contexts, or in cases where the employment of meta-argument allegations becomes common practice for a person (Open-mindedness) or group (Social divisiveness). Others (Violation of norms; Conflict escalation) apply more generally. Together, I submit that they provide reason for a principled wariness in employing meta-argument allegations of offence and harm.

### **3.1 Honesty, accuracy and bias**

A meta-argument allegation can violate basic ethical norms. For example, Andrew might deliberately exaggerate or misleadingly frame Prue's words to bolster his claims of offence and harm. The powerful capacity of meta-argument allegations to demolish constructive argument about politically charged topics allows them to be used strategically, as a deliberate rhetorical tactic to repel a potentially challenging line of argument, or to drag the topic of argument away from an opponent's expertise and experience.

Using a meta-argument allegation in a deliberately misleading way constitutes knowingly saying something false and is a straightforward example of dishonesty. To be sure, deliberate or reckless violation of the principle of charity occurs in ordinary arguments, but in this case it is more pernicious, as: a) it is being used to level an allegation of wrongdoing, and potentially to justify sanctions, b) it is derailing, so exploration of the violation does not shed epistemic light on the issue under consideration, and c) the allegation can have the auto-immunising quality mentioned above (in Sect 2.3), preventing the normal tools of rational argument from uncovering and mitigating the violation.

Even if Andrew's interpretation of Prue's words is not deliberately misleading, it may nevertheless be seriously flawed, as Andrew might be under the sway of hot (emotional) or cold (rational) biases that cause him to badly misinterpret Prue's words (Correia 2012, 227). This would not be surprising. Daniel Kahn (2016, 2-5) notes how intelligent people routinely fall into the trap of politically motivated reasoning, where their critical reasoning faculties are employed not in the service of truth and knowledge, but instead to protect their identities and maintain their in-group status. It is easy to see how the same process could drive the use of meta-argument allegations, with intelligent partisans unconsciously using their rational capabilities to locate sources of offence and harm in their opponents' arguments, and thereby obviating any need to deal with them on their merits. This unconscious instinct may mislead the allegor into avoidable misrepresentations of another's words in the course of levelling a serious moral accusation.

### **3.2 Violation of argumentational norms**

Scholars working on the ethics of argument put forward an array of norms, virtues and principles that they contend are required for respectful argument (Johannesen 1979, 39; Jensen 1981, 20-35; Keller and Brown 1968, 76-79; Ceva 2011, 150; Breakey 2020, 4-5). Commonly defended norms include requirements for openness (that all opposing viewpoints can be aired and must be taken seriously); relevance (discussions should introduce considerations relevant to resolving the issue under consideration) and the foregoing of ulterior tactics to pressure others to agree.

Meta-argument allegations can breach each of these norms. The raising of a meta-argument allegation can breach openness by preventing a newly aired viewpoint being considered on its merits. The allegation can violate relevance by introducing and prioritising an entirely new topic (about offence or harmfulness). And subsequent action based on the allegation (such as deplatforming Prue, pressuring her employment or institutional membership, or hounding her

through attacks in press and social media) can amount to a violation of the prohibition on ulterior pressure.

Why would these breaches of argumentational norms be ethically significant? Scholars put forward a variety of justifications for argumentation norms, including appeals to the requirements of political democracy, social contract arguments and more. To keep things brief, I will note three considerations I have raised in a recent work (Breakey 2020, 13). First, there can be issues of consent in play. Prue may have only begun the interaction on the expectation that Andrew was willing to rationally explore the topic with her, but after her initial, good faith revelation of her convictions, Andrew diverts the topic and shifts from constructive argument with Prue to rhetorical declamations about her. In opening herself to argument, Prue can feel that she has made herself vulnerable, and that Andrew exploited that vulnerability. This concern will be amplified if Andrew begins applying ulterior pressures that move the interaction out of the domain of argument, and into more coercive, “win at all costs” terrain (Crawford 2009, 108; Brockriede 1972).

Second, we noted in Sect 2 that, when conducted well, arguments deliver significant epistemic and cognitive goods. These can link intrinsically and instrumentally with ethical goods like integrity, self-understanding, trustworthiness and moral decision-making. While no-one has an ethical duty to positively contribute to others’ cognitive pursuits, we nevertheless possess duties not to frustrate and derail such pursuits (Breakey 2020, 10). This is especially so for moral perspectives that invoke knowledge as an intrinsic human good that people are called upon to participate in (Finnis 1980, 96). And it is of particular significance from a Kantian perspective, where respect for others’ moral, prudential and theoretical rationality assumes a central place (Formosa 2017, 79). It was on this basis that Kant (1996, 6:463–468) inveighed against treating other’s reasoned arguments with contempt or ridicule, or wanton fault-finding. Meta-argument allegations are concerning on a similar basis to these

wrongs: they can work to stymie, trivialise and demoralise people's efforts at reasoned argument and understanding, and hence to disrespect their rationality.

Third, arguments can contribute to collective decision-making (Breakey 2020, 16-17). In both small and large groups, inclusive deliberations offer a means of arriving at outcomes that all participants can accept. For example, Prue might come to accept a final decision because she has been persuaded by Andrew to change her mind about the issue. Sometimes—even in the most high-stakes and cutthroat contexts imaginable—people do change their mind, even about cornerstones of their positions (Risse 2000, 9, 27). Alternatively, Prue might accept the outcome because the deliberation has led to a new position that she and Andrew both can accept as reasonable (or, at least, as not unreasonable), or perhaps just because she felt she was respectfully listened to and had her views seriously discussed (considerations that fall under the banner of deliberative legitimacy (Breakey 2018, 6-7)). When meta-argument allegations derail such arguments, they stymie this way of dealing with the morally delicate business of collective decision-making.

### **3.3 Conflict escalation**

Meta-argument allegations are a powerful tool for escalating conflicts. This can be seen through Friedrich Glasl's (1999) influential theory of conflict escalation. On Glasl's view, an argument (even a spirited and pointed argument) is not itself a social conflict. To become a social conflict, one agent must see the other as restricting "the way in which she lives out or realizes her own ideas, feelings or intentions" (Glasl 1999, 18). Once this occurs, the opposing parties have a social conflict, and Glasl plots their level of escalation on a 9-point scale, where each level correlates with the parties' increased intransigence, polarisation and hostility. Beginning with incivilities like rhetorical point-scoring (at Level Two), the stages move through to judgements about the opponent's knowledge and abilities (Four), to moral



accusations (Five), threats (Six) and ultimately to “all or nothing” fighting (Eight) until Level Nine sees the parties plunging “together into the abyss”.

There is no inevitability about this escalation. People make choices at each threshold about whether to hold, proceed or attack (Glasl 1999, 106). Yet there are significant social and psychological forces that tend to drive people onwards despite themselves, making retreat and de-escalation difficult.

Glasl’s theory demonstrates the game-changing qualities of meta-argument allegations. He (1999, 75-83) argues that escalation drivers include *snowballing* (the tendency of speakers in a conflict to bring in more issues than the initial one causing the conflict) and *personification* (the tendency to employ “you-language” and to make disparaging claims about people’s character). Meta-argument allegations are a powerful method of snowballing, for the new issues they introduce assume a natural priority (as we saw in Sect.2), making it difficult to get back to the source of conflict. Meta-argument allegations also encourage personification, by introducing assertions of blame and wrongdoing.

In terms of escalation levels, an ordinary argument (that is not even a conflict) rockets straight to Level Five the moment a meta-argument allegation is levelled. In so doing, meta-argument allegations move the disputants into situations where they not only have a conflict, but they have a “conflict about the conflict” (Glasl 1999, 109). That is, they begin to see their opponent’s behaviour in the conflict as itself a matter for conflict.

While some conflicts may be morally necessary, and even desirable, ethical agents generally have good reason to resist avoidable conflict escalation. As escalation occurs, it becomes increasingly difficult for both parties to remain in control of their behaviour. Conflicts have their own logic, and it can require considerable experience, restraint and forbearance to halt the descent, and to scale things back to a civil level.

### 3.4 Open-mindedness

A further area of ethical concern arises if we think that open-mindedness is a moral virtue—either intrinsically, or because it instrumentally helps us escape biased and blinkered views which carry substantial moral risks (such as if they are employed in political decision-making, or to justify intolerance). To be sure, some scholars argue there are limits on the virtue of open-mindedness (Fantl 2018, 27) and that some types of dissent are neither epistemically nor morally valuable (Biddle and Leuschner 2015, 262). Yet in most everyday cases, when diverse views are welcomed and seriously considered in arguments, this helps combat narrowmindedness, bias and groupthink, and encourages more sophisticated and informed epistemic positions (Sunstein 2006, 80; Aikin and Clanton 2010, 410-411; Correia 2012, 236; Ekstrom et al. 2019 17; Green et al. 2000, 1388). Such arguments can be particularly significant in ethical realms. The influential Social Intuitionist Model in social psychology holds that moral reasoning done by individuals on their own tends to be little more than self-interested rationalization of pre-existing intuitions (Haidt and Bjorklund 2006, 181). But moral disagreement *between* people, such as might occur in the course of a principled argument, can empower reason to genuinely impact on subsequent behavior (Haidt and Bjorklund 2006, 181, 190-192).

Unfortunately, Andrew's liberal use of meta-argument allegations may have an insidious impact on his open-mindedness, and his ability to countenance reasoned disagreement. The impact is insidious because Andrew could claim that, *in principle*, he remains open-minded. After all, he is not objecting to the substance of Prue's ideas or arguments, but only to her speech's offensive and harmful parts. But, of course, there is large overlap between the topics and terms that political partisans tend to find offensive and harmful, and the arguments characteristically made by their opponents. This is not a new phenomenon. As Mill (1859/2003, 50-51) observed in the nineteenth century (and as Plato (1997/380BC, 475d) had

perceived more than two millennia earlier), it is easy to feel offence when faced with little more than a strongly argued, contrary position:

I think experience testifies that this offence is given whenever the attack is telling and powerful, and that every opponent who pushes them hard, and whom they find difficult to answer, appears to them, if he shows any strong feeling on the subject, an intemperate opponent.

Consistent use of meta-argument allegations allows Andrew to shut himself off from serious consideration of challenging points, and perhaps even to permanently foreclose entire lines of argument. Worse still, the use of meta-argument allegations could become a cognitive habit. Andrew might start automatically scanning Prue's words for offensiveness or harmfulness—even to the point where he is *expecting* to find such qualities. This would not be a surprising result. We already saw (in Sect. 3.1) how Andrew might be tempted by politically motivated reasoning to dispense with opponents' arguments. He might equally be driven by confirmation bias and selective exposure—the desire to actively avoid information challenging existing views (Hart et al. 2009, 2). There is a real danger that people like Andrew, driven by the powerful instincts behind politically motivated reasoning and selective exposure, could use meta-argument allegations to avoid the difficult and challenging work of seriously engaging with contrary views.

### **3.5 Social divisiveness**

A final line of concern lies in the larger social consequences of widespread use of meta-argument allegations. In an environment where levelling meta-argument allegations is common, when people from different political persuasions argue there will be frequent violations of argumentational norms, regular conflict escalations and increasing opportunities for narrow-mindedness. Instead of people across political divides having the type of positive

interactions that can draw them together (Chua 2018, 201), their attempts at dialogue lead to derailment and outrage. This context will then weaken the prospects for future constructive arguments, as the vulnerability, honesty and reciprocity required (Crawford 2009, 111-118) will be increasingly absent. As Keller and Brown (1968, 75) observe, “our communicative habits, as speakers, are moulded and shaped by the responses we get from listeners.”

In the longer term, differing views on offence and harm, and the eagerness to trade meta-argument allegations upon their basis, can create a disconnect or “empathy wall” (Hochschild 2016, 5) between two sides of politics, where they no longer talk—where they no longer know *how to talk*—to the other side. No side wants to engage with the other when this will just provide their opponents with an opportunity to say offensive and harmful things. Equally, no side wants to broach difficult topics when it is only a matter of time before the argument is utterly derailed, and the topic of conversation once again moves to heated allegations. Once this occurs, groups will tend to talk only to like-minded fellows, contributing to problems of political ‘bubbles’, groupthink and group polarisation, where contrary views are never aired and social dynamics lead to increasingly extreme positions (Peters 2019, 406).

I submit that most people who engage in political argument do not desire this outcome. They do not want a halt to genuine debate between arguers holding different opinions. What they want, I suggest, is for that debate to happen, but to take place on *their* terms, controlled by *their* views of what is offensive, and constrained by *their* views of what is harmful. But once we realize that the very differences that make us disagree on substantive policy will—except for the most obvious and immediate cases—also lead us to disagree on matters of offence and harm, we can appreciate that this wish is fanciful. If we value the capacity, as individuals and as a society, to get together and talk about our disagreements, then we must avoid using those very disagreements as a reason not to talk to each other. But meta-argument allegations do exactly this. They make agreement on fundamental areas of disagreement—on what is

offensive, on what is unhelpful and harmful, and on where responsibility for either lies—a precondition for being able to talk about fundamental areas of disagreement.

## Conclusion

Section 1 acknowledged that epistemic and moral goods can come from meta-argument allegations, and that these will prove decisive when responding to the most egregious cases of harm and offence, such as vitriolic hate-speech or incitement to violence. Outside these cases however, Sections 2 and 3 have offered an array of epistemic and moral concerns that recommend “argumentational tolerance”—a principled wariness in employing meta-argument allegations. This virtue empowers us to respect important norms of ethical argument, value the goods that may be pursued through argument, avoid personal perils of bias and close-mindedness, reduce conflict escalation, and contribute to a larger culture that is supportive of rational, inclusive debate. Argumentational tolerance thus warrants inclusion within existing lists of argumentational virtues and ethics (Thorson 2016, 363-365; Aberdein 2010, 171-175; Correia 2012, 231-237).

Notwithstanding these moral reasons for embracing argumentational tolerance, the foregoing arguments do not justify morally “calling out” a meta-argument allegation whenever it is levelled. For one thing, there will be times (like hate speech) when meta-argument allegations are manifestly legitimate. But for another, the reality is that levelling a *meta*-meta-argument allegation is likely to trigger all the worrying costs noted above, including the qualities of derailment, violation of argumentational norms and conflict escalation. Instead, the foregoing argument provides resources for an interlocuter like Prue to try to help Andrew see that the different perspectives that make their argument worthwhile are bound to provide them with different views about the nature and significance of offence, harm, benefit, risk and responsibility. Prue is no more likely to agree to constrain her arguments, and tiptoe around

the expression of her ideas, on the basis of Andrew's particular views, than Andrew is to make the same unilateral concession to Prue. To be sure, it may be that Andrew and Prue come in time to converge on many of these issues. But such convergence is only likely to happen as a *result* of their constructive argument. For this reason, they both have reason to avoid making their substantive disagreements a barrier to that very argument.

**Acknowledgements** This paper was presented at the 2017 Australian Association for Professional and Applied Ethics Annual Conference and in the 2018 University of Queensland School of Historical and Philosophical Inquiry Research Seminar Series. I am grateful to discussants at both events, and also to Paul Formosa and two anonymous reviewers for *Argumentation* for their insightful comments on a previous draft.

## References

- Aberdein, Andrew. 2010. Virtue in Argument. *Argumentation* 24:165-179.
- Aikin, Scott F., and J. Caleb Clanton. 2010. Developing Group-Deliberative Virtues. *Journal of Applied Philosophy* 27 (4):409-424.
- Biddle, Justin B., and Anna Leuschner. 2015. Climate skepticism and the manufacture of doubt: Can dissent in science be epistemically detrimental? *European Journal for Philosophy of Science* 5:261–278.
- Blair, J. Anthony. 2012. Rhetoric, Dialectic, and Logic as Related to Argument. *Philosophy and Rhetoric* 45 (2):148-164.
- Breakey, Hugh. 2018. It's right, it fits, we debated, we decided, I agree, it's ours, and it works: The gathering confluence of human rights legitimacy. *Law and Philosophy* 37 (1):1-28.
- Breakey, Hugh. 2020. The Ethics of Arguing. *Inquiry* Forthcoming.
- Brockriede, Wayne. 1972. Arguers as Lovers. *Philosophy & Rhetoric* 5 (1):1-11.
- Ceva, Emanuela. 2011. Just interactions in value conflicts: The Adversary Argumentation Principle. *Politics, Philosophy & Economics* 11 (2):149-170.
- Chua, Amy. 2018. *Political Tribes: Group Instinct and the Fate of Nations*. London: Bloomsbury.
- Chung, Sungeun, and Shin-Il Moon. 2016. Is the Third-Person Effect Real? A Critical Examination of Rationales, Testing Methods, and Previous Findings of the Third-Person Effect on Censorship Attitudes. *Human Communication Research* 42 312–337.
- Cohen, Daniel. 2001. Evaluating arguments and making meta-arguments. *Informal Logic* 21 (2):73-84.
- Cohen, Daniel. 2009. Keeping an Open Mind and Having a Sense of Proportion as Virtues in Argumentation. *Cogency* 1 (2):49-64.

- Coppola, Vincent, and Fabien Girandola. 2018. When Increasing the Strength of the Argument Becomes Counterproductive: The Role of Argumentative Markers in the Arousal of the Psychological Reactance. *Journal of Language and Social Psychology* 37 (4):475–496.
- Correia, Vasco. 2012. The Ethics of Argumentation. *Informal Logic* 32 (2):222-241.
- Crawford, Neta. 2009. Homo Politicus and Argument (Nearly) All the Way Down: Persuasion in Politics. *Perspectives on Politics* 7 (1):103-124.
- Cross, Ben, and Louise Richardson-Self. 2019. “Offensiphobia” is a Red Herring: On the Problem of Censorship and Academic Freedom. *The Journal of Ethics*:1-24.
- Davison, W. Phillips. 1983. The Third-Person Effect in Communication. *The Public Opinion Quarterly* 47 (1 ):1-15.
- Eemeren, Frans van, Bart Garssen, and Bert Meuffels. 2012. The disguised abusive ad hominem empirically investigated: Strategic manoeuvring with direct personal attacks. *Thinking & Reasoning* 18 (3):344-364.
- Eemeren, Frans van, and Rob Grootendorst. 2003. *A Systematic Theory of Argumentation*. Cambridge: Cambridge University Press.
- Ekstrom, Pierce D., Brianna A. Smith, Allison L. Williams, and Hannah Kim. 2019 Social Network Disagreement and Reasoned Candidate Preferences. *American Politics Research* (1–23).
- Fantl, Jeremy. 2018. *The Limitations of the Open Mind*. Oxford: Oxford University Press.
- Finnis, John. 1980. *Natural Law and Natural Rights*. Oxford: Clarendon.
- Finocchiaro, Maurice A. 2013. *Meta-argumentation: An Approach to Logic and Argumentation Theory*. London: College Publications.
- Fogelin, Robert, and Walter Sinnott-Armstrong. 2001. *Understanding Arguments: An Introduction to Informal Logic*. 6th Aufl. Orlando: Harcourt College Publishers.
- Formosa, Paul. 2017. *Kantian Ethics, Dignity and Perfection*. Cambridge: Cambridge University Press.
- Gampa, A., S. Wojcik, M. Motyl, B. A. Nosek, and P. Ditto. 2019. (Ideo)logical reasoning: ideology impairs sound reasoning. . DOI: 10.1177/1948550619829059. *Soc. Psychol. Pers. Sci.*:1-9.
- Glasl, Friedrich. 1999. *Confronting Conflict*. Gloucestershire: Hawthorn.
- Goldberg, Michelle. 2014. Feminism’s Toxic Twitter Wars. *The Nation* February:12-17.
- Green, M. C., P. S. Visser, and P. E. Tetlock. 2000. Coping with accountability cross-pressures: Low-effort evasive tactics and high-effort quests for complex compromises. *Personality and Social Psychology Bulletin* 26:1380-1391.
- Haidt, Jonathan. 2012. *The Righteous Mind: Why Good People are Divided by Politics and Religion*. London: Penguin.
- Haidt, Jonathan, and Fredrik Bjorklund. 2006. Social Intuitionists Answer Six Questions About Moral Psychology. In *Moral psychology, Vol. 2. The cognitive science of morality: Intuition and diversity*, ed. W. Sinnott-Armstrong, 181–217. London: MIT Press.
- Hart, William, Dolores Albarracín, Alice H. Eagly, Inge Brechan, Matthew J. Lindberg, and Lisa Merrill. 2009. Feeling Validated Versus Being Correct: A Meta-Analysis of Selective Exposure to Information. *Psychol Bull.* 135 (4):555-588.
- Hirschman, Albert. 1991. *The rhetoric of reaction: Perversity, futility, jeopardy*. Cambridge: Harvard University Press.
- Hochschild, Arlie Russell. 2016. *Strangers in Their Own Land: Anger and Mourning on the American Right*. New York: New Press.
- Jensen, J. Vernon. 1981. *Argumentation: Reasoning in Communication*. New York: Litton Educational Publishing.

- Johannesen, Richard. 1971. The emerging concept of communication as dialogue. *The Quarterly Journal of Speech* 57 (4):373-382.
- Johannesen, Richard. 1979. Perspectives on Ethics in Persuasion In *Persuasion: Reception and Responsibility*, ed. Charles Larson, 26-42. Belmont, Cal.: Wadsworth.
- Kahan, Dan M. 2016. The Politically Motivated Reasoning Paradigm, Part 2: Unanswered Questions. In *Emerging Trends in the Social and Behavioral Sciences*, eds. Robert Scott, and Stephen Kosslyn. New York: John Wiley & Sons.
- Kant, Immanuel. 1996. The Metaphysics of Morals. In *Practical Philosophy*, ed. Mary J. Gregor. Cambridge: Cambridge University Press.
- Karp, Paul. The Guardian (Australia Edition). 2017. Homophobic anti-marriage equality material surfaces in postal survey campaign. 21 August 2017.
- Keller, Paul, and Charles Brown. 1968. An Interpersonal Ethic for Communication. *The Journal of Communication* 18:73-81.
- Kim, Sang-Yeon, Timothy R. Levine, and Mike Allen. 2017. The Intertwined Model of Reactance for Resistance and Persuasive Boomerang. *Communication Research* 44 (7):931–951.
- Kriesberg, Louis. 2007. *Constructive Conflicts: From Escalation to Resolution*. 3rd Aufl. Oxford: Rowan and Littlefield.
- Lewis, Rosie. The Australian. 2016. Gay marriage plebiscite could lead to suicide: Shorten. September 12 2016.
- Maitra, Ishani, and Mary Kate McGowan. 2012. Introduction and Overview. In *Speech and Harm: Controversies Over Free Speech*, eds. Ishani Maitra, and Mary Kate McGowan. Oxford: Oxford University Press.
- Mill, John Stuart. 1859/2003. *On Liberty*. Cambridge: Hackett.
- Peters, Uwe. 2019. Implicit bias, Ideological Bias, and Epistemic Risks in Philosophy. *Mind & Language* 34:393–419.
- Plato. 1997/380BC. Gorgias. In *Plato: Complete Works*, ed. John M. Cooper, 791-869. Cambridge: Hackett.
- Risse, Thomas. 2000. “Let’s Argue!” Communicative Action in World Politics. *International Organization* 54 (1):1-39.
- Shorten, Bill. 2016. Marriage Legislation Amendment Bill 2016: Second Reading. 12 September 2016, ed. Parliament of Australia. Canberra: Australian Government Hansard.
- Sunstein, Cass. 2006. *Infotopia*. New York: OUP.
- Thorson, Juli K. 2016. Thick, Thin, and Becoming a Virtuous Arguer. *Topoi* 35:359-366.
- Tirrell, Lynne. 2012. Genocidal Language Games. In *Speech and Harm: Controversies Over Free Speech*, eds. Ishani Maitra, and Mary Kate McGowan, 174-221. Oxford: Oxford University Press.
- UN General Assembly. 1965. International Covenant on the Elimination of All Forms of Racial Discrimination. New York: United Nations.
- Vavova, Katia. 2018. Irrelevant Influences. *Philosophy and Phenomenological Research* XCVI (1):134-152.
- Walton, Douglas N. 1989. Dialogue Theory for Critical Thinking. *Argumentation* 3:169-184.