



Is Rational Manipulation Permissible?

Hugh Breakey¹ 

Accepted: 13 March 2023
© The Author(s) 2023

Abstract

Rational manipulation is constituted by the following conditions: (i) A aims to persuade B of thesis X; (ii) A holds X to be true and rationally justifiable; (iii) A knows of the existence of evidence, argument or information Y. While Y is not itself misinformation (Y is factually correct), A suspects B might take Y as important evidence for not-X; (iv) A deliberately chooses not to mention Y to B, out of a concern that it could mislead B into believing not-X; and, (v) B has no compelling reason to expect A will avoid mentioning Y in this way. A's behavior is *rational* insofar as A aims to use reasons to persuade B to believe a thesis that A holds as true and justified. Yet it is *manipulation* because A deliberately avoids furnishing B with information that B might regard as relevant, to ensure B arrives at the correct belief. I argue that we have good reason to think that A's action will be wrongly manipulative because it disrespects B's consent, epistemic autonomy, and personal autonomy. That said, context is critical, and there are many times evidence is intentionally occluded that are not rational manipulation. Even so, ethical arguers should beware of ambiguous contexts, such as when there are conflicting expectations about argumentation roles and goals.

Keywords Autonomy · Ethics of argument · Epistemic paternalism · Science communication · Manipulation

1 Introduction

Imagine Anna and Bob are discussing the ethics of abortion. Anna knows that there are certain facts about fetal development that, if Bob knew, he would likely take as reason to be wary about Anna's strong pro-choice position. However, Anna believes Bob would be mistaken in arriving at this conclusion—she thinks the facts in question, while true, would

✉ Hugh Breakey
h.breakey@griffith.edu.au

¹ Institute for Ethics, Governance & Law, Griffith University, Brisbane, Australia

mislead Bob and draw him into error. On this basis, Anna deliberately avoids making Bob aware of those facts in their discussion.

Has Anna acted wrongly? Does Bob have any legitimate grievance if he later realizes that Anna treated him this way? It is easy to think that he does not. After all, if we consider the matter from Anna's perspective, during the argument she is providing accurate and relevant reasons supporting what she believes is a true conclusion in a case where her interlocuter, Bob, has consensually engaged with her to discuss the issue. What—she might reasonably demand—could possibly be wrong with that?

In answer, I argue that Anna is guilty of *rationaly manipulating* Bob, and that there are good reasons for thinking she thereby wrongs Bob by disrespecting his autonomy. However, context matters. I explore cases where similar behavior would not be wrongful, as well as vexing cases where the status of rational manipulation remains ambivalent. These arise acutely in cases where Anna has ambiguous role-expectations.

The paper proceeds as follows. After this introductory section, Section 2 defines rational manipulation. Section 3 provides three autonomy-based reasons for thinking that rational manipulation is wrongful. Section 4 repels three potential counter-arguments. Section 5 considers several vexing argumentation contexts, giving rise to challenges and ambiguities.

Ultimately, I argue that rational manipulation is a serious matter, and one to which ethical arguers must attend. At the same time, context is crucial, and there are times when the deliberate suppression of known evidence can be justified.

2 Rational Manipulation Defined

Rational manipulation is constituted by the following necessary and sufficient conditions:

1. *Persuasive aim*: A aims to persuade B of thesis X;
2. *Sincere thesis*: A holds X to be true and rationally justifiable;
3. *Misleading evidence*: A knows of the existence of evidence, argument or information Y. While Y is not itself misinformation (Y is factually correct), A suspects B might take Y as important evidence for not-X;
4. *Deliberate suppression*: A deliberately chooses not to mention Y to B, out of a concern that it could mislead B into believing not-X;
5. *Innocent expectations*: B has no compelling reason to expect A will avoid mentioning Y in this way.

To keep the exposition clear, I will adopt the following terminology. I will refer to A as the *persuader* and B as the *interlocuter*. I will refer to the information (Y) the persuader elects to suppress as the *evidence* (noting that Y may be a line of argument, a new concept, information, data, or any other knowledge the persuader possesses that the interlocuter can be expected to find relevant).

As a definitional matter, there is good *prima facie* reason for thinking the persuader's behavior is rightly termed "rational manipulation" (Tsai 2014: 90). The persuader's behavior is *rational* in two senses. First, the persuader aims to rationally persuade the interlocuter to believe the thesis by providing him with sound argumentation: true premises that link together logically to imply the conclusion. Second, the persuader believes that her thesis

is both true and rationally justified: she thinks that an informed and conscientious agent, reasoning logically on the basis of all available evidence, would accept it.

At the same time, the persuader's behavior is *manipulative*. Standard definitions of 'manipulation' are, inter alia, (a) to control or play upon by artful, unfair, or insidious means especially to one's own advantage; and (b) to manage or utilize skillfully (Mirriam-Webster 2023). The persuader's behavior is manipulative in both senses. The persuader controls the interlocuter's thinking in an insidious way, because her action occurs, and is intended to occur, beneath the interlocuter's conscious awareness. But also, the persuader's control is done artfully and skillfully, not by lying crudely to her interlocuter, but by keenly anticipating his likely reasoning and curating his knowledge. She plays him like a puppet.

This understanding of manipulation coheres with Ana Nettel and Georges Roque (2012: 58) describing manipulation in the context of persuasion as an intentional strategy that involves *dissimulation*—especially about the persuader's true agenda and priorities, and the means they will use to pursue these—and *constraint*—which can involve not fully presenting all relevant reasons and alternatives (see also Wilkinson 2013: 347). Nettel and Roque acknowledge that both manipulation and argumentation can create persuasion, but argue that argumentation's capacity to persuade is not ethically concerning because by providing reasonable arguments the persuasion happens with the interlocuter's acceptance and consent. While I think that we should broadly accept Nettel and Roque's analysis of both manipulation and argument, I will argue that rational manipulation presents a special case where the providing of reasonable arguments (in concert with intentionally suppressing relevant evidence) can count as ethically wrongful manipulation. Argument *can* manipulate.

Let us now consider severally rational manipulation's five constitutive conditions. In exploring these conditions, I do not imply that argumentation that avoids one of these five conditions is necessarily morally permissible. Such argumentation might well be morally wrong, perhaps even wrong for the same autonomy-based reasons (see Sect. 3) which make rational manipulation wrong. But, as I hope to show, rational manipulation presents specific ethical concerns which warrant careful analysis.

2.1 Persuasive Aim

Rational manipulation only occurs in context of argumentation, when the persuader is actively trying to make the case for a specific thesis to the interlocuter. This need not require disagreement. The persuader might aim to simply maintain, bulwark, or strengthen an existing belief the interlocuter holds (Aikin and Casey 2022). Indeed, rational manipulation typically will be easier in cases without disagreement, because the discussion will appear cooperative and non-adversarial (see 'Innocent Expectations' below). Yet rational manipulation still has serious effects in these cases, including by contributing to group polarization, rationalization, and radicalization through systematically suppressing contrary considerations (Aikin and Casey 2022: 134).

In contrast, imagine in our initial scenario that Anna suppresses information she thinks Bob will mistakenly find relevant—but simply because she wants a high-quality reasoned discussion, and doesn't want to waste time dealing with red herrings. Anna is not curating Bob's informational environment so she can win, and any charge of *manipulation* therefore looks inapt. Anna may contribute to Bob holding a particular belief, but this was not her intention.

2.1.1 Sincere Thesis

The persuader is not leading the interlocuter to a conclusion she knows to be false. She is therefore not straightforwardly misleading or gaslighting him.

2.1.2 Misleading Evidence

The persuader is conscious of the evidence, believes it to be true, and is aware of its potential relevance for the interlocuter. However, the persuader has no positive duty to scour their mind for such evidence. It is because the persuader knows the information, acknowledges it is true, and anticipates its relevance for the interlocuter, that the concern with manipulation arises.

2.1.3 Deliberate Suppression

Rational manipulation requires the *intentional suppression* of evidence. If the evidence is mentioned, but only in the “fine print” or as an offhand remark in another context, then—if this is done for purposes of suppression—this is rational manipulation. However, foregrounding and prioritizing other evidence, and framing the argument to focus on points that the persuader takes to be more probative, do not count as rational manipulation.

This is important, as people never have unlimited time or space to present their arguments. It may be that in their available time, the persuader has only enough room to cover what she sees as the most important arguments for her thesis, and not enough time to mention potential countervailing concerns, or to explain why she thinks these concerns are irrelevant.

In such cases, the persuader has a strong response against charges of manipulation. She can agree that if they had time, it would have been right for her to share evidence she knew the interlocuter would find relevant. However, the persuader eschewed relating such evidence because she wanted to cover what she saw as the most telling points.

Still, this excuse only goes so far. In the cases of detailed (or repeat) communiques and lengthy discussions, there will come a point where the persuader will have the opportunity to address the evidence. If the persuader fails to do so out of a concern that the interlocuter will be misled, rational manipulation occurs.

2.1.4 Innocent Expectations

The *Innocent Expectations* condition requires that the interlocuter has no compelling reason to expect the persuader will suppress Y. If the interlocuter knows the persuader is going to suppress known evidence then any charge of manipulation is harder to sustain, because the interlocuter should know to *caveat emptor*. He can make a decision about whether engaging with the persuader is worthwhile to him under these conditions, and consent (or not) as he likes. Because he knows he is not getting the full story, he can seek out other sources of information.

The interlocuter’s expectations can be shaped in four main ways.

First, there may be an explicit declaration that certain types of information will be deliberately suppressed. For example, there are strict rules on the types of information—such as

“similar fact evidence” about a defendant’s past misconduct—that can be presented to jurors (Goldman 1991; Ahlstrom Vij 2013). Jurors are informed about these constraints, so when lawyers later present arguments that intentionally suppress information, the jurors are aware that this might be happening.

Second, B. J. Diggs (1964: 367-9) observes that in certain cases we can have positive obligations to persuade. Politicians advocate for political standpoints and constituencies. Salespeople are employed to sell products. While normal ethical rules still apply to these advocates (e.g., they should not lie), they are expected to provide only evidence and arguments supporting their persuasive goal.

Third, some types of arguments are highly and self-evidently adversarial. In the “Dominant Adversarial Model”, each side does its best to win the argument by successfully persuading the other and/or an audience (Stevens 2019). An even more adversarial dialogue is the *quarrel*, where the opponent’s defeat or humiliation is the over-riding goal (Walton 1998). In these cases, the arguer’s behavior will typically signal that she is adopting an unapologetically adversarial role and style of argument (Stevens 2019), allowing the interlocuter (inter alia) to expect the persuader will suppress contrary evidence.

Stevens and Cohen (2021) helpfully distinguish having an adversarial *attitude* from other types of adversarial qualities and roles—such as aiming to defend a thesis, and to show the weaknesses of opposing arguments. The overriding goal of an arguer with an adversarial attitude is to *win*. This might be done through aggressive and intimidating behaviour. In this case, exemplified in the quarrel, rational manipulation cannot occur, because the arguer’s manifestly cutthroat behaviour shows that the interlocuter can hardly expect positive assistance. However, Stevens and Cohen (2021: 901) acknowledge that the goal of winning might be pursued in other, less aggressive ways, “intended to prevent reasons which might work against their goals from being recognized.” Rational manipulation is such a case, and it is one where a disjuncture arises between an arguer’s *professed role*, which may appear cooperative, and their *actual intentions*, which are to ensure at all costs that they achieve their persuasive goals. Rational manipulation can only occur when an arguer’s adversarial attitude is to some extent hidden—the attempt to bludgeon an interlocuter into defeat conflicts with seducing them through manipulation (Brockriede 1972).

These points about adversarial arguments and advocacy combine with the earlier Persuasive Aim condition to constrain the types of dialogues where rational manipulation can arise. As we saw above, rational manipulation can occur in many different cases of argument—including cases where there is little or no disagreement between the arguers (Aikin and Casey 2022). But it can also occur in persuasion dialogues—or at least dialogues containing discrete instances of attempted persuasion (Walton 1989, 1998). At the same time, it cannot be a context where the persuaders’ oppositional role as a negotiator, advocate, or adversary is *so pronounced* that the interlocuter should expect the suppression of contrary evidence. While this might make rational manipulation in persuasion dialogues seem only narrowly applicable, these conditions remain relatively common because—as discussed below—persuaders are more persuasive when they present themselves as honest brokers.

Fourth, there may be cases where the persuader has a known expertise and set of values that provide the interlocuter with a clear understanding of areas they might resist discussing. For example, if the persuader is a scientist or science teacher, then the interlocuter could hardly expect the persuader to bring up pseudo-scientific or conspiratorial information (Goldman 1991: 121). The scientist can reasonably insist that if the interlocuter had

concerns on those bases, then they should have known the scientist would be an inappropriate discussant.

These four ways of shaping expectation cabin the extent of rational manipulation. We routinely expect others to give less than frank arguments, and in such cases rational manipulation does not occur. However, the word *compelling* in the Innocent Expectations condition is deliberate. As discussed below, when expectations are murky or conflicting, rational manipulation can occur. The persuader might think it is self-evident that she has an advocacy role, or is arguing in an adversarial context. But if the interlocuter has reason to think differently, then they may be wrongfully manipulated.

2.1.5 Rational Manipulation's Relationship to Epistemic Paternalism

Epistemic paternalism refers to paternalistically curating a subject's knowledge environment to improve their epistemic outcomes. Alvin Goldman (1991: 119) invoked epistemic paternalism as occurring any time information controllers interposed their own judgment to improve their audience's epistemic prospects, rather than allowing the audience to exercise their own judgment. Kristoffer Ahlstrom Vj (2013: 39–51) similarly described epistemic paternalism as interfering with an agent going about inquiry as they see fit, without consulting them, and for their epistemic benefit.

On either understanding, rational manipulation emerges as a specific case of epistemic paternalism, as rational manipulation involves non-consensual interference to improve epistemic outcomes. However, not all epistemic paternalist activities involve rational manipulation. Epistemic paternalism can occur even if it is expected, such as if we are aware we are being epistemically nudged. (Some nudges arguably manipulate (see Wilkinson 2013), but if done transparently they are not rational manipulation.) Epistemic paternalism might be done to improve overall epistemic environments, without intending to persuade—focusing on *how* people reason, rather than *what* they should believe.¹ Finally, epistemic paternalism need not involve suppressing truth. It can involve suppressing *untruths*, or requiring people to speak frankly (Goldman 1991: 122; Godden 2021), or simply involve framing and the prominent placement of probative evidence. Such framing might still be wrongfully paternalistic (Tsai 2014), but it is not rational manipulation.

Summing up, even if rational manipulation is wrong, other types of epistemic paternalism might be justified. That said, below I critique epistemic paternalist arguments where they might seem to justify rational manipulation.

3 Is Rational Manipulation Wrongful?

Rational manipulation wrongfully undermines autonomy. I argue there are three relevant ethical principles of autonomy that rational manipulation transgresses: the principle of consent (did the interlocuter agree?); the principle of epistemic autonomy (does this respect the

¹ Ahlstrom Vj (2013: 95) asserts optimistically that, “Epistemically paternalistic interventions are not designed to tell people *what* to believe, but *how* to come to believe things.” But this constraint is not reflected even in his own definition (2013: 39), meaning epistemic paternalism can be done to persuade, and therefore can (at least potentially) rationally manipulate.

interlocuter as a thinker and knower?); and, the principle of personal autonomy (does this respect the interlocuter’s capacity to govern their life?).

3.1 Autonomy and Consent

When we interact with people—especially when we aim to influence or change them—we should do so on the basis of their consent, or at least in ways to which we expect they would consent. In Kantian terms, if two people interact on the basis of expected, known standards, and then one departs from the standards for some discretionary end the other doesn’t share, then this can constitute the first person treating the second person as a mere means, and not as an end-in-themselves (Kant 2008: ¶428, Formosa 2017: 92). These concerns apply to argument. Richard Johannesen (1979: 27) highlights the ethical concerns when a persuader’s persuasive strategies violate an “implied agreement” with their interlocuter (see also Breakey 2020: 13).

Rational manipulation violates this standard. The persuader deliberately suppresses information that she thinks the interlocuter would find relevant to his judgment, and does not alert him to this fact. Plausibly—absent special circumstances—the interlocuter cannot be assumed to have consented to the persuader’s strategy. After all, it is natural for the interlocuter to want to make their decision about the thesis on the basis of all the evidence that they would judge as relevant, and it is precisely this evidence that the persuader is suppressing.

Why would the rational manipulator avoid alerting the interlocuter to her strategy? Because it makes her case more persuasive. As Diggs (1964: 364) reflects,

When one not only offers advice or recommends, but also tries to persuade, he is giving a “strong,” “thorough,” or “complete” kind of advice or recommendation; he is attempting to do more. . . . One who does more has a greater responsibility. When one attempts to persuade, he presumes to tell another what he should believe or how he ought to act: he is not just offering help; he, so to speak, has made another’s decision of what to believe or do for him, and is trying to get the other to accept the decision.

In other words, persuaders in argument typically present themselves as giving a thorough or complete case. After all, their case is more compelling if they present themselves as being fully informed, frank and honest brokers on the matter. It is in this gap between what they are presenting and what they are suppressing where the interlocuter’s consent is being disrespected.

Simply put, rational manipulation fails a “sunlight test”. It is a practice that, to be effective, cannot be done with the interlocuter’s understanding and consent.

3.2 Epistemic Autonomy

The second autonomy-related concern enjoins us to specifically respect others’ *reasoning faculties*. Others, like us, are capable of rational thought. They can think things through for themselves. They can respond to reasons—meaning that if we want to change their mind, we can do so by supplying reasons and evidence, and eschew resorting to more coercive or underhand methods. In addition, we can learn from them, acknowledging that they have

important and worthwhile experiences, evidence, ideas, and lines of thought. Further, we can think things through with them—by engaging in deliberation we can come into agreements and shared cognitive states (Cohen and Miller 2016). Acknowledging our interlocutor as rational in all these ways, and treating them on this basis, establishes an important equality between us, and is a significant way we can show ethical respect (Kant 1996: ¶6:463–468; Formosa 2017: 79). Ultimately, people are both somewhat rational and somewhat irrational, and we respect their epistemic autonomy when we choose to appeal to and engage with the former capability, rather than exploiting the latter vulnerability. The point of epistemic autonomy, so understood, isn't to ignore others' testimony, arguments, communications, or behaviour (see Ahlstrom Vaj 2013: 93–95), but to critically *weigh these up* for ourselves, such as by asking whether we have reason to trust them, and if they make sense to us.

The rational manipulator disrespects the interlocutor's epistemic autonomy. Deciding that they know better, the persuader intentionally chooses not to supply the interlocutor with all the information and let them make up their own mind. Instead, they make allowances for the interlocutor's presumptively poor rationality by usurping its working and using their own rationality—beneath the interlocutor's conscious awareness—to lead them to the persuader's preferred judgement. This violates a key tenet of ethical argument—the crucial acceptance that others are ultimately entitled to come to their own conclusions (Breakey 2020: 4).

In so doing, the persuader commits a moral wrong by undermining or ignoring the interlocutor's status as a *knower*. Miranda Fricker (2007) uses the term “epistemic injustice” to refer to a class of wrongs where a person is discriminated against in a way that—on the basis of unfair prejudice against the person's group (e.g., gender, race)—undermines or ignores their status as a knower (Fricker 2007). Plausibly, it is this larger political dimension (prejudice or wrongdoing against a politically marginalized group) that warrants Fricker's use of the term “injustice” (see also Fricker 2013). Rational manipulation can be an epistemic injustice in this sense. On the basis of the interlocutor's already marginalized identity, the persuader might prejudicially decide to forego reasoned persuasion and engage in rational manipulation. The persuader would thus wrong a person in their capacity as a knower in a way that both stems from, and can feed into, an existing socio-political marginalization. Evan Riley (2017) puts forward a version of epistemic injustice termed *reflective incapacitatorial injustice*, that occurs in wrongful failures to support those in marginalized groups in the development and/or exercise of their reflective capacities for critical reasoning. Rational manipulation can do exactly this; it aims to suppress evidence to avoid engaging with a perceived weakness in the interlocutor's thinking, rather than providing the evidence and debating its probative value. When a marginalized person is rationally manipulated in this way, this therefore constitutes an epistemic injustice—specifically, a reflective incapacitatorial injustice.

However, many cases of rational manipulation will not have this political dimension. They are an instance of what we might simply term an “epistemic immorality”—wrongfully undermining or ignoring an individual's status as a knower qua individual person (rather than qua member of a politically marginalized group). Plausibly, many types of ethical wrongdoing in argumentation contexts (e.g. Tsai 2014; Stevens 2019; Breakey 2020) will count as epistemic immoralities in this sense.

Rational manipulation not only fails to engage with the person qua rational agent, it also has potentially concerning epistemic outcomes. Granted, successful manipulation will

ensure the interlocuter comes to a true belief they might otherwise not have had. This is an epistemic benefit. But (subject to potential exceptions discussed in Sect. 5 below) that belief is not resiliently well-supported by the interlocuter's reasons. The interlocuter does not understand the full case for or against their belief, and therefore the belief is vulnerable to the subsequent appearance of the suppressed evidence. As well, the rational flaws or blind spots that made the manipulation necessary have not been addressed. The interlocuter has lost the opportunity to be mistaken about the thesis, and learn from their mistake. Indeed, as Stephen John (2018: 82) observes, the suppression of evidence can even strengthen the pre-existing rational flaw (such as a naively idealistic view of science)—making the same mistake *more* likely in future. Worse still, if ultimately exposed, the suppression will likely cause the interlocuter resentment, making him less likely to trust the (perhaps otherwise epistemically beneficial) persuader. Expressed in David Godden's (2021) apt terms, rational manipulation might improve the interlocuter's epistemic *situation*, even as it worsens their epistemic *agency*.

Summing up, respect for another person's epistemic autonomy involves appealing to, engaging with, and supporting their capacities for reasoned thought. Rational manipulation violates this required respect.

3.3 Personal Autonomy

Respecting others' epistemic autonomy (as above) urges us to allow each person to come to their own opinion because they, no less than we, have rational faculties, and they may possess knowledge and experiences that we should not too quickly discount, dismiss or override. Respecting others' personal autonomy focuses not on the *quality* of others' rational faculties, but on the appropriate *role* of those faculties. On this understanding of autonomy as self-determination, the proper ethical role of each person's reasoning faculty is to govern their life—to allow each person to “shift for themselves” as Locke expressed it (1988: ¶II:60, 83). Simply, even if you have demonstrably superior reasoning powers to me, that does not give you the right to rule over my thoughts, life, or actions, because your rational faculties are properly directed to running *your* life, not *mine*. On this footing, as Joseph Raz (1986: 377–8) argues, a person is autonomous if they determine the course of their life by themselves—making coercion and manipulation signature ethical wrongs.

Is this type of autonomy relevant to rational manipulation? Godden (2021) argues that, because all humans must be committed to the norms of belief, epistemic autonomy is best understood in terms of self-governance through acting on the basis of rules (i.e. *epistemic* autonomy), rather than exercising self-determining choices (i.e. *personal* autonomy). However, facts, values, risk-appetites, responsibilities, attributions of trustworthiness, dialogic goals, and all-things-considered practical judgments entangle in complex ways (Goldenberg 2016; Furman 2020; Walton 1998) that make it difficult to cleanly distinguish epistemic autonomy from personal autonomy. Yet even if Godden is correct in his articulation of epistemic autonomy, it does not follow that interferences in people's epistemic activities cannot wrongfully threaten personal autonomy. There are two ways this can occur.

First, if the persuader successfully manipulates the interlocuter's beliefs, then this weakens the autonomous self-determining quality of the interlocuter's derivative choices about actions and plans. This looks morally concerning, because, as George Tsai (2014: 85) observes, “in certain deliberative situations, what matters is not simply making the

‘right’ choice, but having one’s choice count as ‘one’s own.’” Instead of leaving it up to the interlocuter to make up their own mind, the persuader chooses to usurp the interlocuter’s judgment. Such wrongdoing is evocatively captured by Wayne Brockriede’s (1972: 5) analogy between sex and argument, and the specter of the *seducer*, who—in aiming to win by beguilement—“tries to eliminate or limit his co-arguer’s most distinctively human power, the right to choose with an understanding of the consequences and implications of available options.”

Second, making choices about learning is a key part of personal autonomy. Choosing when, what, and how to learn—and who to learn with—provides an important way that people take responsibility, constitute themselves, explore their individuality, and define who they are, making the capability to make choices in this domain a vital human freedom worthy of ethical respect and non-interference (Breakey 2012: Ch.5). As Locke (2008: 456) famously declared: “he is certainly the most subjected, the most enslaved, who is so in his understanding.” Locke is hardly alone in this view. A striking array of political theorists across the political spectrum—from Marx to John Stuart Mill, from John Dewey to Gadamer—have argued that engaging in self-directed learning about facts and values is an intrinsically and even quintessentially valuable human capability (Breakey 2012: 117–118; see also Riley 2017: 604, 608).

In contrast, Ahlstrom Vij (2013) argues that personal autonomy provides no defense against epistemic paternalist interference. He cites Raz’s (1986: 422) view that, “paternalism affecting matters which are regarded by all as of merely instrumental value does not interfere with autonomy”. Ahlstrom Vij (2013: 81) then argues that the types of areas where epistemic paternalism would be considered are precisely those where the goods being sought are instrumentally valuable. For scientists wanting to gauge the safety of pharmaceuticals, or jurors getting correct verdicts on a defendant’s guilt, the process of belief formation, and the belief itself, matter precisely because of the subsequent good achieved—they are *instrumental means* to valuable ends.

But this misunderstands Raz’s argument. A good can be instrumentally valuable to one person even as it is intrinsically valuable to another. If we are concerned with autonomy, then the question is whether the good is instrumentally or intrinsically valued *by the subject herself*. (No doubt the rational manipulator sees the interlocuter’s belief, and belief formation processes, in instrumental terms. This instrumentalism is precisely the worry: that they are treating people’s judgments and faculties as instruments instead of intrinsically valuable ends.) So the question becomes: Do people value the ability to think things through for themselves, to inquire in their own way, and to make up their own minds? Or do they value true beliefs only for their instrumental utility in subsequent endeavors? As soon as the question is correctly posed, the answer is plain: people do value the ability to think and inquire for themselves—and (on the two bases of personal autonomy noted above) many philosophers have agreed with them.

Perhaps this does not rule out some of the very specific epistemic interventions Ahlstrom Vij considers. There are good reasons for thinking that pharmaceutical scientists would take public safety, and jurors would take justice for defendants, as of over-riding importance. In these special cases, decision-makers just want to get the answer right, and would expect and agree to constraints serving that purpose.²

² Ahlstrom Vij (2013: 46–47) classes these cases as paternalistic, because, he contends, clinical researchers were not *consulted* on the Food and Drug Administration’s (FDA) demand for randomized controlled

3.4 The Threat of Mistaken and Biased Rational Manipulation

Even an infallible epistemic agent can fail to respect an interlocuter's consent, rationality, and autonomy. However, additional concerns arise for fallible humans. Rational manipulation occurs when the persuader holds the thesis to be true and rationally justifiable, but that is not the same as the thesis *being* true and rationally justifiable. If the central exculpatory factor for the manipulation (the failure to respect the interlocuter's consent, epistemic autonomy, and personal autonomy) lies in the claim that the interlocuter will end up epistemically better off (Goldman 1991; John 2018: 82), then the persuader needs to be extremely confident that this will indeed occur. Unfortunately, people planning to interfere with others' beliefs are beset by the same cognitive biases as those they manipulate. In particular, humans tend to systematically overplay their own rationality and the veracity of their beliefs (Ahlstrom-Vij 2013: 16), making them think they are epistemic authorities, and that others are easily duped or manipulated, meaning people will be cognitively biased to think that their interferences in others' thinking will be justified.

Furthermore, by not being fully frank with the interlocuter about all the reasons for and against their thesis, the persuader helps quarantine their belief in the thesis from the interlocuter's rational interrogation. In so doing, the persuader does not allow that the interlocuter's thinking might inform their own. This is a lost epistemic opportunity—and one that is emblematic of the persuader's disrespect of the interlocuter's rationality.

3.5 Political Concerns

Suppose for a moment that the rational manipulator is not a person but a state authority. In this scenario, a new set of ethical concerns arise—rendered additionally serious by the agents' greater power and scope of control (Goldman 1991: 127). First, there is a concern that a system (a set of laws, or communication control practices) put in place to suppress misinformation (i.e., false evidence) might slip into secretly suppressing misleading but true evidence—and so commit rational manipulation. Second, a system that suppresses misleading but true information might further slip into suppressing politically undesirable but true information: outright political censorship.

Twitter's response to Covid misinformation arguably provides a chilling example of both concerns. The release of internal company emails showed that, alongside policing misinformation, Twitter—at the behest of multiple US government agencies—secretly suppressed truthful (and even scientifically supported) concerns about the efficacy and effects of vaccinations and government pandemic policies (Zweig 2022). If the concerns marshalled above are correct, then efforts at correcting misinformation must be handled extremely carefully, lest rational manipulation trigger legitimate resentment and counter-productively undermine vital trust in mainstream epistemic authorities (see Furman 2020).

trials. But paternalism ethically hinges on *consent*. Plainly, when voluntarily choosing to perform research to inform FDA decision-making, clinical researchers know and consent to FDA requirements. FDA requirements are therefore not rational manipulation. They are not even paternalism.

4 Three Counter-Arguments Considered

This section considers three arguments against judging rational manipulation as morally impermissible.

4.1 Rhetoric isn't Special

Back in the early twentieth century, responding to what he felt were overly wrought ethical challenges to rhetoric, William Schrier (1930) observed that there were many non-rational ways of persuading and influencing people. These have only grown more common since Schrier wrote. They include catchy jingles, shiny graphics, using makeup and perfume, employing attractive models in advertising, eye-catching roadside signs, and so on (see also Godden 2021). If we judge these attempts at non-rational persuasion as morally acceptable, then it appears inconsistent to object to similar types of non-rational persuasion (like flattery) that occur in the context of rhetoric.

Like the non-rational persuasive devices just noted, rational manipulation works beneath conscious rational processes. However, rational manipulation nevertheless has features that make it more worrying than these quotidian influencers. First, rational manipulation works by infiltrating our rational process itself. The interlocuter can't take a deep breath and step back to think about things more objectively, because it is precisely in this domain that they have been manipulated. Second, rational manipulation occurs in cases where it is not expected and anticipated. Non-rational influences are much less concerning when they can be anticipated, consented to, or avoided.

4.2 Belief is the Interlocuter's Responsibility

As Diggs (1964: 364) rightly observed, interlocuters have responsibilities too. They are not merely passive subjects being influenced, but rational agents capable of taking responsibility for their own beliefs. After all, there is nothing preventing the interlocuter from further pursuing their thinking in this area and getting more information elsewhere. Can it therefore be argued it is the interlocuter's responsibility not to fall for the persuader's tricks?

In response, the interlocuter doesn't know what he doesn't know. As we saw above, Diggs (1964: 364) himself observed that persuaders—in order to persuade—typically present themselves as trustworthy, well-informed, and constructive, leaving the interlocuter with no reason to think that they are only getting one side of the story. Just as lying is morally wrong—even though a victim could assiduously search out the truth—so too is rational manipulation. Note though that, in a case where the persuader elects not to mention evidence she knows *the interlocuter already knows*, this is not plausibly manipulation, as in this case the interlocuter can reasonably shoulder the responsibility of introducing the known evidence into discussion.

4.3 The Countervailing Importance of true Beliefs

Even though they are manipulating the interlocuter, the persuader might feel they are justified given the moral importance of the interlocuter holding true beliefs.

First, the persuader might justify the manipulation on the basis of the importance of truth. After all, virtuous arguers are ones who “propagate truth” and “spread true beliefs around” (Aberdein 2010: 173, see also Aikin and Clanton 2010). Normally, these are morally worthy properties. They might gain moral worth from taking knowledge to have intrinsic value (Finnis 1980), or perhaps based on a version of the “ethics of belief” (Clifford 1877). More broadly, truth is useful to have, compared to error, and the interlocuter is more likely to make good prudential and moral decisions if they hold true beliefs. While the persuader would need to be wary of undermining longer term epistemic goods (as noted above), they may nevertheless judge their suppression to be in the interlocuter’s best epistemic interests.

Second, the persuader might manipulate the interlocuter to altruistically improve the world. Arguments—and who wins them—have effects in the world, and sometimes these possess profound moral importance. Giving help to your interlocuter (and even more to your opponent) by alerting them to contrary evidence risks frustrating the objective of successfully changing their mind (or, perhaps, an audience member’s mind) and the good outcomes that would result from this.

How compelling are these consequences in justifying rational manipulation? Perhaps in especially rare, urgent, and high stakes circumstances these consequences become definitive (just as we might consider *in extremis* violating one innocent person’s rights to avert a catastrophe). But in our ordinary moral lives, good consequences are not considered sufficient reason to violate an individual’s consent, epistemic autonomy, and personal autonomy, and there is no reason why rational manipulation should be treated differently.

5 Cases Where the Moral Wrongs of Rational Manipulation May be Mitigated

5.1 Monological Versus Dialogical Argument

I have been speaking as if the persuader and interlocuter are engaged in an argumentative exchange, a *dialogue*. But argument can take a monological form, such as if Anna writes a treatise to a general audience, aiming to persuade readers like Bob to adopt her thesis. In such cases, rational manipulation may still occur. This could happen if Anna is aware that many readers—perhaps her *typical* readers—would find certain evidence germane, but she elects not to share it out of concern that they will thereby fail to accept her thesis. Still, Anna can hardly accommodate *every* reader’s pet concerns, given she is writing for a wide audience, and that many other readers might find her excursions irrelevant and distracting. In such cases, she has a legitimate reason for electing not to relate all the evidence she knows.

5.2 Ambiguities in Expectations

We observed in Sect. 2 that expectations are critical, and that in many cases interlocuters have compelling reason to expect persuaders might be suppressing relevant evidence. Unfortunately, there are many ambiguous cases. As Douglas Walton (1989: 175) observed, when arguers have different understandings of the type of dialogue they are mutually engaged in, ethical concerns can quickly arise. Rational manipulation provides one example of how this can occur.

Consider op-ed writers and opinion leaders on traditional or social media that may have a clear and unapologetic ideological bent. Unlike official advocates, they are not obliged to defend a given thesis, and fair-minded commentators can and do engage with arguments that cogently challenge their position. In this case, the expectations are murkier, and—I submit—rational manipulation remains possible.

A medical doctor might also find themselves in an ambiguous situation. Operating on the basis of their patient's best medical interests, they might aim to persuade their patient to accept a particular therapy. As an expert in medical science, they can be reasonably expected to focus on scientific matters, and suppress other information. But they are also bound by their professional ethics to get their patient's *informed* consent (Beauchamp and Childress 2009)—meaning their patient might expect the doctor to raise any information that the doctor believes the patient would find relevant to their decision-making. These conflicting expectations might make rational manipulation possible.

Finally, personal relationships may complicate matters. If you need a new car, and your best friend is a car salesperson, then you might think this fact will trump their ordinary role-obligations, and that they will tell you frankly about any information they know you would find relevant to your purchasing decision.

In short, the natural complexity and ambiguity of everyday life provides many cases where expectations are ambiguous, and rational manipulation becomes possible.

5.3 Patent Irrationality and Known Irrationalities

Respecting epistemic autonomy implies a substantive standard. Rationality is both a value (something that people can care about and take seriously, or spurn and trivialize) and it is an epistemic standard (implicating logic, evidence, truth, resistance to bias, etc.). Treating someone as if they had poor or flawed rational faculties—when there is no evidence this is so—is clearly an epistemic immorality. But what about cases where the persuader *does* have substantial evidence that people—or even just this particular person, on this one particular matter—is demonstrably irrational? Consider an obvious case: a parent teaching a child, or a primary school teacher teaching children. Decisions about what not to teach the child will often be made on the basis of the effective use of time. Teaching children about flat earthers' beliefs might require a lot of effort to debunk. Yet there surely will be cases where decisions about exclusions are made precisely because the evidence might lead to wrongful belief (Goldman 1991: 121). In many such cases, there might be a mutually acknowledged epistemic authority that leads to the interlocuter's expectation (and perhaps implicit or explicit consent) that they will be guided by the persuader. The interlocuter acknowledges: "There are things you understand that I don't yet understand, and I consent to trusting your decisions on what to teach and what not to teach." Even so, the expectations and consent here can be murky. A child who becomes an adult can still, in at least some cases, be resentful about not being informed about some true evidence that they (would then, and do now) consider relevant.

Presuming that expectations and consent are not definitive, then these types of teaching decisions will be strictly speaking cases of rational manipulation. However, they are not wrongful on two grounds.

First, and most straightforwardly, the ethical need for respecting autonomy is far weaker in children, meaning the reasons for worrying about the rational manipulation's wrongdo-

ing are mitigated, and often entirely absent. The term “rational curation” is perhaps more appropriate in these cases.

Second, the teaching of children should aim over the long term to develop their epistemic agency. The child should eventually be in a position to decide for themselves about the teaching’s appropriateness. I mentioned earlier that Godden (2021) observes it is possible to improve a person’s epistemic *situation*, while worsening their epistemic *agency*. He suggests a justificatory condition for epistemic paternalism based on this distinction, such that an epistemic paternalist interference should rationally empower the subject to the point where they can subsequently understand and endorse that interference. While one-off manipulations seem to me to be unlikely to achieve such empowerment, the long-term curation of children’s learning may well do so, and therefore be justified in Godden’s sense.

While these two points sensibly accommodate adult-child interactions, there are harder cases of known irrationality and cognitive blind-spots in ordinary adults. In the context of science communication about climate change, John (2018) argues that openness about the workings of scientific systems should not be an ethical norm. As the “climategate” controversy showed, transparency can undermine people’s credence in important truths like global warming. This is because, John argues, laypeople possess a false “folk philosophy of science” (2018: 81). Practices of dogmatism and the selective exclusion of data sets are (John contends) proper parts of the scientific method—but these clash with the idealized folk philosophy. This means that when laypeople are provided with evidence of such “normal and respectable” practices, they mistakenly question the scientists’ claims (John 2018: 81).

John urges we therefore have ethical reason to deliberately hide the system’s workings, on the basis that, “experts’ communicative obligations towards non-experts should, ultimately, be grounded in claims about what will further non-experts’ epistemic interests” (2018: 81). Note how far this differs from ordinary people’s communicative obligations—such as not to lie—which are bound by obligations to respect others as such (including by respecting their consent, epistemic autonomy, and personal autonomy). Note also that there is no analogy here to professional obligations, such as those of a lawyer, that can supersede laypeople’s normal obligations. Professional obligations are public knowledge, they are often enshrined in democratically mandated law, and they are surrounded by complex governance systems to manage conflicting ethical concerns (e.g., between confidentiality and disclosure). In contrast, John’s argument calls for experts’ system-wide rational manipulation: deliberately and secretly suppressing true facts for beneficial epistemic ends.

John provides no argument about why ordinary obligations, rooted in basic respect, no longer apply, except for a fleeting suggestion that the “regrettable” deceitfulness “may be mitigated” by its epistemic outcomes (2018: 82). But this is precisely to treat others as means and not ends, to see their beliefs as instruments to be manipulated to desirable ends, rather than as parts of a person that are owed intrinsic respect. John (2018: 85) later invokes exactly this concern in intrinsically prohibiting deception for *non-epistemic* ends. But epistemic ends have no special status that make deception and manipulation permissible. To the contrary, by deliberately manipulating the very seat of the person’s cognitive capabilities, they are *additionally* ethically concerning.

This does not mean that scientists should include mention of every concern (or at least popular concerns) people may have with their work—including pseudoscientific and other concerns. As described earlier, a scientist’s expertise is in science, and she can be expected to focus on scientific evidence and methods. But it does mean we should listen carefully

to people's concerns, rather than quickly characterizing them as unreasonable, irrational, or unscientific (Goldenberg 2016; Furman 2020). After all, as noted above, even if people plainly suffer from various irrationalities and cognitive flaws, we can still make a decision whether to use these limitations as the definitive basis for our engagement with them (and so rationally manipulate them), or whether we can instead focus on where and how their reasoning is working well, and the sensible and understandable concerns they might have that we could work to address, if only we listened carefully enough to what they were saying (see Furman 2020, and, especially, Goldenberg 2016).

To be sure, it is possible to imagine an interlocuter who struggles to distinguish between science and pseudoscience, between evidence and fancy. This interlocuter does have reason (objectively speaking) to expect a scientist to confine her arguments to scientifically respectable considerations, but that reason is not personally accessible to him. Because the interlocuter can't distinguish science from pseudoscience, he may be frustrated and resentful when the scientist constantly avoids talking about issues the interlocuter sees as relevant. He is not being rationally manipulated—but he might think he is. This result suggests that people with very poor belief revision processes may tend to think everyone is (and all mainstream institutions are) constantly trying to manipulate them, by persistently not discussing information they think is manifestly relevant.

6 Conclusion

Rational manipulation presents an intriguing topic in the ethics of argument. It involves an array of important moral concerns—respect for consent, respect for epistemic autonomy, and respect for personal autonomy. At the same time, the way those concerns play out in each situation are subtly but importantly influenced by myriad contextual factors: adversarial settings, logistical constraints, advocacy roles, and the persuader's knowledge and intentions. For all these reasons, this is not an area where definitive moral judgments can be easily delivered. Yet the stakes in play are high, and instances of rational manipulation—as epistemic injustices or epistemic immoralities—may provoke resentment, anger, and mistrust. Virtuous arguers have reason to keep expectations clear, to listen hard, and to err on the side of frankness and respect for their interlocuters' rational faculties.

Acknowledgements This article has benefited from critical comments from Katharina Stevens, Andrew Aberdein, three anonymous reviewers, and from participants' comments at the Ethics of Argumentation Speaker Series (July 1, 2023).

Funding Open Access funding enabled and organized by CAUL and its Member Institutions

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

References

- Aberdein A (2010) Virtue in Argument. *Argumentation* 24:165–179. <https://doi.org/10.1007/s10503-009-9160-0>
- Ahlstrom-Vij K (2013) *Epistemic paternalism: a defence*. Palgrave Macmillan, New York
- Aikin S, Casey J (2022) Argumentation and the problem of agreement. *Synthese* 200(2):134. <https://doi.org/10.1007/s11229-022-03680-4>
- Aikin SF, Clanton JC (2010) Developing Group-Deliberative Virtues. *J Appl Philos* 27(4):409–424. <https://doi.org/10.1111/j.1468-5930.2010.00494.x>
- Beauchamp T, Childress J (2009) *Principles of biomedical ethics*. Oxford University Press, New York
- Breakey H (2012) *Intellectual Liberty*. Routledge, London
- Breakey H (2020) The Ethics of Arguing. *Inquiry*. <https://doi.org/10.1080/0020174X.2019.1637776>
- Brockriede W (1972) Arguers as lovers. *Philos Rhetoric* 5(1):1–11
- Clifford W (1877) /1999 The ethics of belief. In: Madigan T (ed) *The ethics of belief and other essays*, Amherst, MA, Prometheus pp 70–96
- Cohen D, Miller G (2016) What Virtue Argumentation Theory misses: the case of compathetic argumentation. *Topoi* 35:451–460. <https://doi.org/10.1007/s11245-015-9334-3>
- Diggs BJ (1964) Persuasion and ethics. *Q J Speech* 50(4):359–373
- Finnis J (1980) *Natural Law and Natural Rights*. Clarendon, Oxford
- Formosa P (2017) *Kantian Ethics, dignity and perfection*. Cambridge University Press, Cambridge
- Fricker M (2007) *Epistemic injustice: power and the Ethics of Knowing*. Oxford University Press, Oxford
- Fricker M (2013) Epistemic justice as a condition of political freedom? *Synthese* 190. 1317–1332. <https://doi.org/10.1007/s11229-012-0227-3>. 7
- Furman K (2020) On trusting neighbors more than experts: an Ebola Case Study. *Front Communication* 5(23):1–8. <https://doi.org/10.3389/fcomm.2020.00023>
- Godden D (2021) Epistemic autonomy, epistemic paternalism, and blindspots of reason. In: Axtell G, Bernal A (eds) *Epistemic paternalism: conceptions, justifications, and implications*. Rowman & Littlefield, London, pp 181–197
- Goldenberg MJ (2016) Public misunderstanding of Science? Reframing the Problem of Vaccine Hesitancy. *Perspect Sci* 24(5):552–581. https://doi.org/10.1162/POSC_a_00223
- Goldman A (1991) Epistemic paternalism: Communication Control in Law and Society. *J Philos* 88(3):113–131
- Johannesen R (1979) Perspectives on Ethics. In: C Larson (ed) *Persuasion: Reception and Responsibility*. Belmont, Cal., Wadsworth: pp 26–42
- John S (2018) Epistemic trust and the ethics of science communication: against transparency, openness, sincerity and honesty. *Social Epistemology* 32(2):75–87. <https://doi.org/10.1080/02691728.2017.141086>
- Kant I (1996) *The Metaphysics of Morals*. Practical philosophy. M. J. Gregor. Cambridge University Press, Cambridge
- Kant I (2008) *The moral law: Groundwork of the Metaphysic of Morals*. Hutchinson's University Library, New York
- Locke J (1988) *Two treatises of government*. Cambridge University Press, Cambridge
- Locke J (2008) *An essay concerning Human understanding*, 2nd edn. Thomas Basset
- Miriam-Webster (2023) *Manipulation*. Miriam-Webster Dictionary
- Nettel AL, Roque G (2012) Persuasive argumentation Versus Manipulation. *Argumentation* 26:55–69. <https://doi.org/10.1007/s10503-011-9241-8>
- Raz J (1986) *The morality of freedom*. Clarendon, Oxford
- Riley E (2017) The beneficent Nudge Program and Epistemic Injustice. *Ethical Theory and Moral Practice* 20(3):597–616. <https://doi.org/10.1007/s10677-017-9805-2>
- Schrier W (1930) The ethics of persuasion (a defense of rhetoric). *Q J Speech* 16(4):476–486
- Stevens K (2019) The Roles we make others take: thoughts on the Ethics of Arguing. *Topoi* 38:693–709. <https://doi.org/10.1007/s11245-019-09659-0>
- Stevens K, Cohen D (2021) Angelic Devil's advocates and the Forms of Adversariality. *Topoi* 40(5):899–912. <https://doi.org/10.1007/s11245-020-09726-x>
- Tsai G (2014) Rational persuasion as Paternalism. *Philos Public Affairs* 42(1):78–112. <https://doi.org/10.1111/papa.12026>
- Walton D (1989) Dialogue theory for critical thinking. *Argumentation* 3:169–184
- Walton D (1998) *The New Dialectic: conversational contexts of argument*. University of Toronto, London
- Wilkinson TM (2013) Nudging and manipulation. *Polit Stud* 61(2):341–355. <https://doi.org/10.1111/j.1467-9248.2012.00974.x>

Zweig D (2022) How Twitter Rigged the Covid Debate. The Free Press. December 27, 2022. <https://www.thefp.com/p/how-twitter-rigged-the-covid-debate>

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.