

## 4 Don't Block the Exits

*Justin Tosi*<sup>1</sup> and *Brandon Warmke*<sup>2</sup>

John Stuart Mill famously argued in *On Liberty* that freedom could be threatened not only by the state, but by society in general. Mill was especially concerned about the effects of social coercion on free expression. For even if the state is legally constrained from interfering with individual expression, that is little consolation to those who are fired from their jobs or ostracized by their communities for expressing their beliefs.

We share Mill's concern, which is sadly only becoming more pressing. According to one recent survey, 62 percent of Americans are afraid to share some of their beliefs because of the current political climate.<sup>3</sup>

Evidently, our purportedly liberal society has not embraced Mill's warnings about social coercion. Americans enjoy robust First Amendment protections against state interference with free expression. Yet most of us feel highly vulnerable to interference from everyone else.

What would a society look like that more closely realized Mill's vision for a free liberal order? Such a society, we think, would have very different social norms about matters of free expression than we currently have. Some of these more liberal norms are easy to imagine. For instance, a truly liberal society would have a norm against trying to get someone fired just because you disagree with her political views. But other norms would likely be less straightforward, as some problems of social coercion arise because of complex considerations of incentives and human psychology.

This essay considers the phenomenon of what we call *exit-blocking*. Roughly speaking, someone engages in exit-blocking when she imposes a social cost on another person for changing her mind. We'll explain this phenomenon and consider some likely motivations that drive people to behave this way. We'll then argue that exit-blocking is both a moral and epistemic problem. Finally, we discuss some norms against exit-blocking that a society might adopt to discourage it and thereby enhance the freedom of its members.

## 4.1 What Is Exit Blocking?

In a society characterized by ideal liberal freedom, people would be free to form their beliefs solely on the basis of epistemic considerations and not because of social pressure. The same would be true of deciding whether to speak their minds in public.

Our own society does not live up to this ideal. One of the reasons it falls short is that some people apply social pressure in ways that discourage others from changing their minds about moral or political matters. Or at least, some people apply social pressure in ways that discourage people from *talking in public* about changing their minds. In other words, some people block the exits and thus make others less free to form their beliefs solely on the basis of reasons and evidence.

Of course, not everything you might do to influence another's beliefs or public statements will count as exit-blocking. Disappointed to see that someone you thought of as a fellow progressive has changed her mind about an important issue, you might present her with objections or counterarguments. Or you might show her social scientific evidence that the causal claims her reasoning relies on are implausible. Your interventions might sway her to revise or abandon her new position. Alternatively, she might regard your second-guessing as tiresome or annoying. In any case, your behavior as described exerts influence of a permissible kind, because it consists solely of epistemic pressure. Even if a person doesn't want to confront possible flaws in their beliefs, pointing them out doesn't impose the kind of cost we're concerned with.

When a person blocks the exits, he does so by imposing a social cost on another person for changing her mind. By social cost, loosely speaking, we mean some social result of changing your mind that you might dread. Paradigmatic examples include humiliation in front of your community, loss of relationships, or, in the most extreme case, ostracism.<sup>4</sup> A social cost is, in other words, a negative impact on your social world—your reputation, relationships, associations, and so on.

Not all instances of the phenomenon are highly costly for the target. Publicly issuing personal insults can be a form of exit-blocking, for example. Although most people are not deeply bothered by such responses, it isn't pleasant to be on the receiving end of them, and so the threat of being insulted can deter rational people from speaking up.

Our attention will mainly be devoted to forms of exit-blocking that are less costly, since it is harder to see what is so troubling about them for their own sake, and so easier to focus on the broader dynamics we are concerned about. We will also talk mostly about exit-blocking in public discourse, though it could, in principle, also happen in private exchanges.

Now some sensitive souls might insist that being subjected to *any* public rebuttals of their claims should be treated alongside these cases as

social costs. And to be fair, it can indeed be embarrassing for someone to give an objection to your view that you can't answer, and not everyone enjoys arguing about ideas. So we agree that being called to defend your views is a kind of cost, but it is not the kind we are concerned about here.

The kinds of social costs we have in mind are those that result from what Linda Radzik calls "informal social punishment" (Radzik 2020). Informal social punishment intentionally imposes a harm on someone for allegedly violating a norm. Examples of such informal social punishment include: calling someone a fascist in front of her colleagues, emailing her boss to try to get her fired, posting her email address online to incite harassment, and so on. Unlike *formal* social punishments—like getting disinvited from a conference, getting fired, or getting banned from Twitter—informal punishments are not mediated by institutions or institutional procedures.

Exit-blockers therefore try to impose a harm—understood in terms of a social cost—on those who change their minds about moral and political issues, which the exit blocker considers a norm violation. However, simply providing an objection or rebuttal to someone who has changed her mind will not typically be a form of informal social punishment, even if the person changing her mind bears some social cost, such as being embarrassed for having her views challenged in public.

People block the exits for many different reasons, and their behavior takes diverse forms. There is no single motivation behind exit blocking. And there's no single kind of exit-blocking behavior. A helpful starting point for mapping the conceptual terrain is to consider the group dynamics of this phenomenon. Here is one simple way of thinking about the group dynamics of exit-blocking: upon publicly changing your mind, you can be challenged either by members of your in-group or out-group. In-group exit-blocking happens when members of your in-group impose social costs on you for changing your mind. Out-group exit blocking happens when your out-group imposes social costs on you for changing your mind.

In-groups and out-groups will typically have different incentives to challenge reported changes of mind, and their exit-blocking behavior will often look different. Yet both in-group and out-group exit blocking comes to the same thing: discouraging others from changing their beliefs.

From one standpoint, it might seem like members of an out-group should regard it as a welcome event when someone changes his mind to agree with them. But sometimes out-group members do not meet such conversions with approval. Instead, they attack these newly like-minded people, and so block their exit. In such cases, out-group members will say things like, "this change of heart is too little too late." Often, they criticize the convert's timing or the slow pace he took to arrive at the right view, saying, for example, "if you couldn't see that this is a deadly

disease until now, you're an idiot, and you should really sit this one out." Or, "where were you four years ago when we could have actually used your help to keep a fascist out of the White House? Sorry, not buying it." Also common are baseless accusations that newcomers to the cause are changing their minds now only because it is convenient to do so, or because they finally want to be on the right side of history. Sometimes people mock in advance those who might be tempted to change their minds, predicting that history's losers will come crawling to the enlightened group once they realize their children despise them, and we should never let them forget their mistakes. Out-groups block the exits from the outside.

In contrast, in-group members block the exits from the inside, imposing social costs on apostates who publicly disavow their previously shared beliefs. One common approach is to use the target's *other beliefs* to accuse them of being inconsistent, or more to the point, of standing for nothing on principle. Someone blocking an exit from the inside might say, "How could you now think this election was legitimate when you've been railing against these cheaters for years like the rest of us?" We can easily imagine someone saying this as a request for clarification or to reconcile two beliefs that are apparently in tension. But the intention of the exit-blocker in cases like this is to attack the moral purity or integrity of the apostate. She thinks this person has betrayed the cause, and wants to plant the idea in others' heads that the person is insincere or hasn't really thought the matter through, because he is corrupt or not even trying to get it right.

As we have said, exit-blocking can take many forms. In-group and out-group members can say similar things to block the exits, and they can even literally make the same statements or ask the same questions. Still, exit-blocking statements by the in-group and out-group are distinguished by the position they put the target in relative to that group. The in-group accuses the target of never truly believing in the first place or of being corrupted. The out-group insists he doesn't truly believe now or has changed his views for the wrong reasons. Both faces of exit-blocking are often reducible to purity tests. What they have in common even more fundamentally is the imposition of a cost on the target, and a deterrent effect on those who might follow suit.

Why does anyone engage in exit-blocking? Some cases can probably be attributed to straightforward malice. Some exit blockers probably just want to harm others and see an opportunity to do so.

But we think there are additional explanations for the phenomenon. One is that when a person changes his mind about some issue of interest to a group, it creates an opportunity for members of that in-group and its related out-groups either to seek more social status, or to project and solidify their status in their group. Psychologists have identified two forms of status that people seek, prestige and dominance. Prestige status

is conferred to a person because of her admirable qualities, skills, or achievements. Dominance status is earned by being seen as someone with control over things people care about, or by being intimidating or threatening (Cheng et al. 2010).

We suspect that a lot of exit-blocking is a form of moral grandstanding (Tosi and Warmke 2016, 2020). When people grandstand, they use public moral discourse to seek social status. Grandstanding members of a target's in-group might see her apostasy as an opportunity to make a morally flattering contrast with themselves. They capitalize on this opportunity by seeking prestige status. Whereas the apostate doubts or rejects the correct moral views, we have much purer beliefs. She is abandoning the right side of history while we remain firmly on it, as always.

Grandstanding out-group members, on the other hand, might see recent converts as a threat to their status. Why should this person be praised for finally waking up when we have been right all along? The correct views have always been obvious to us, and this dunce has only just now figured it out. Or, she may have gotten this one right, but she's still wrong about everything else, so we can't trust her. Grandstanders from both groups see that their moral qualities can be favorably compared to people who have recently changed their minds, so they do so in order to gain or maintain prestige.

Those who change their beliefs are also ripe targets for grandstanding to seek dominance status. When you change your mind publicly, you weaken the incentive for those in your in-group to defend you, or to be seen associating with you. Even if you convert entirely to the beliefs of what was previously an out-group, the members of that group might not trust you or value your association with them yet. Thus, by changing your mind, you present aggressive people with a chance to exercise their will to power. They can lash out at you to demonstrate their capacity for aggression without fear of reprisal from your tribe.<sup>5</sup>

Another way of understanding exit-blocking is to see it as an expression of what psychologists call the black sheep effect (Marques et al. 1988; Pinto et al. 2010). The black sheep effect is the phenomenon of deviant or weak members of a group being evaluated more harshly than similar members of out-groups. Weak members of a group are less committed to the norms and beliefs of the group than strong members. They are thus viewed as unreliable, or untrustworthy. Those who announce that they have changed their mind about something important to their in-group are, in a sense, willingly becoming marginal members of their group. The stronger members downgrade their evaluation as a result, and they may respond by announcing their disappointment.

A useful extension of this paradigm is that it helps explain the appeal of exit-blocking as a strategy for deterring further defection from the group. The individual apostate is the occasion for the exit-blocking, but others in the group who may be teetering on the edge of changing their

own minds can see the response and recognize the incentive not to do so. Nobody wants to be the black sheep.

## 4.2 Moral and Epistemic Criticisms of Exit-Blocking

From what we have said so far, it should be clear that people typically don't enjoy being on the receiving end of exit-blocking. But exit-blocking isn't just a distasteful experience. It is a bad social practice, for both moral and epistemic reasons. It makes us less free, and it interferes with truth-seeking.

Exit-blocking is morally bad because it interferes with freedom of expression. Many people are understandably reluctant to air their beliefs in public when they run counter to the orthodoxy within their community. This reluctance will only be stronger when a community engages in exit-blocking, as it raises the cost of changing your mind. People in such communities must worry not only about the silent judgment of their peers, but also about public campaigns expressing that disapproval when they stray from the flock.

Mill understood that informal social pressure could interfere with our freedom just as state coercion does. In fact, tyranny by an informal collective is, Mill says, "more formidable than many kinds of political oppression, since... it leaves fewer means of escape, penetrating much more deeply into the details of life, and enslaving the soul itself" (1989, 8). Informal communities are not bound by legal rules and procedures in the way state actors are. Any group member can simply smear an apostate for being morally or politically impure because of her public comments, and it becomes a liability for anyone else in the group to be seen associating with that person. The community might not be able to imprison or put to death its apostates, but it can make their lives miserable in ways states cannot. It can deny them personal relationships, willing trading partners, or any of the other benefits of social life, just by discouraging others from associating with them lest they suffer the same fate.

A prudent person would thus try to avoid the costs of social disapproval for changing her mind in matters that might draw exit-blocking responses. If she is not so cautious, the penalties could be severe. Members of their in-group will be wary of them, since the exit-blockers have marked her as impure, not to be trusted. Out-group members will also see her as impure, since she has been suspiciously slow to come around to the right view. The result is that freedom is limited to picking a team quickly and restricting one's public statements to claims that cohere with the group's views.

Mill argued that freedom of expression is "practically inseparable" from freedom of conscience (1989, 15). To limit expression is therefore to impinge upon the most intimate region of human liberty. If a society

regards it as normal and acceptable behavior to impose social costs on those who change their minds, then it encourages its members either to maintain a private conscience that does not leak into the public sphere, or to extinguish their heterodox thoughts before they might lead to public apostasy. Many will find it too mentally taxing to do the former, so they will instead bring their private mental life into line with the views acceptable within their group. Their freedom is thus limited.

But we should not restrict our evaluation of exit-blocking to its moral costs. Exit-blocking exerts a pernicious influence not just over what people are free to say, but what they believe. The incentives established by exit-blocking motivate us to change the way we think about the topics our groups care deeply about. As a result, people are discouraged from changing their minds, even privately.

In a world of ideally free inquiry, the social cost of changing one's mind would be low. People would be able to consider arguments and evidence without being pressured to come to a particular conclusion. Thus, they would be free to believe for epistemic reasons alone. Exit-blocking interferes with free inquiry. When social costs for changing one's mind are imposed, people are disincentivized to take opposing views seriously, seek out countervailing evidence, or even to find out what the opposition thinks. If you are already attached to some view, and your social or professional relationships would be damaged should you abandon that view, then you have strong incentives not to entertain doubts about it. Exit-blocking is thus an invitation to engage in motivated reasoning. Knowing that you will face some social cost for changing your mind, you might quite understandably look for ways to avoid doing so, even if only subconsciously. You might be more easily swayed by *prima facie* plausible but fallacious arguments for your view, or by strawman characterizations of opposing views, for example.

It might be objected that our concern about people being worse at even private reasoning in the face of increased social costs is overblown. Surely we can still evaluate views on their merits when we are just thinking about them on our own, and not report our conclusions to others if we think they won't like what they hear. It is true, of course, that most of us don't report our every thought publicly, and that among the thoughts we are most likely to keep to ourselves are the ones that will draw social blowback. But our point is not that people can't reason privately and safely become heterodox thinkers that way. Rather, it is that, at the margins, fewer people will do so. It is unpleasant to live an intellectual double life, always worried that you will accidentally say something that doesn't fit with the web of socially acceptable beliefs. In addition, people do not want to think themselves cowardly. But if you develop a set of beliefs that you cannot reveal to others without paying a high social penalty, it is hard not to realize that you *are* afraid to stand up for what you believe in. Instead of landing in such an unpalatable position, many

people will find some way to retain socially acceptable beliefs and just not think about them too much.

Exit-blocking also influences our private beliefs by limiting the arguments and evidence that people hear in public discourse. Hrishikesh Joshi has helpfully analyzed this dynamic as a collective action problem (2021, 67–85). It is important that people be exposed to evidence they are not aware of, or else we will collectively suffer from significant blind spots. We can alleviate those blind spots by sharing our evidence, and thereby exposing people to considerations that they would not have discovered on their own. But it is also risky for any individual to speak her mind and present evidence that might not please the crowd. If the risks of speaking one's mind are great enough, we might all quite rationally decide that it isn't worth it to speak up.

When groups engage in exit-blocking, they raise the risk—and thus the cost—of speaking your mind. Rational actors see that there is little to gain by expressing beliefs that their in-group might not approve of. Their own social network will label them traitors, and out-groups will refuse to offer them safe harbor. And worst of all, even if they did accept the cost of speaking out, their individual sacrifice would probably produce little overall benefit. The rational thing to do, then, is to hope other people discover the same thoughts as you, and are less risk-averse about sharing them. But the greater the risk, the fewer such people there will be.

The result of this collective action problem is that people are left without the social resources that might help them reason their way out of false beliefs. This might be good for group cohesion, but it is bad for free inquiry. Where exit-blocking is prevalent, loyalty is valued more than truth-tracking.

### **4.3 Toward an Environment Safe for Free Inquiry**

Having described the phenomenon of exit-blocking and explained why it is bad for both moral and epistemic reasons, we'll now present a vision of a society that doesn't suffer from this problem. We will describe some potential norms that we could adopt to ensure greater overall freedom than we currently enjoy in our present society of exit-blocking enthusiasts.

We will follow Cristina Bicchieri in using “social norm” to refer to “a rule of behavior such that individuals prefer to conform to it on condition that they believe that (a) most people in their reference network conform to it (empirical expectation), and (b) that most people in their reference network believe they ought to conform to it (normative expectation)” (2016, 35). An individual's reference network is the group of people whose behavior and attitudes are relevant to them for social purposes. In the case of these norms, a person's reference network will likely include members of both their in-group and out-groups. The in-group



is relevant because these are the people we rely on most for our important social relationships, and we want to remain in good standing with them by following their norms. Out-groups are relevant primarily out of considerations of reciprocity: if they are not following these norms, then other groups may feel less bound by them.

The most obvious norm that should be adopted is against exit-blocking itself. We should not criticize people or impose other social costs on them simply for changing their minds. Again, this does not mean we should refrain from criticizing their justifications for doing so. But shaming people as traitors or apostates, or for not seeing the light sooner, is a bad practice.

A related and supporting norm is that we should not reward exit-blocking behavior. People respond to incentives, and if they can gain status for insulting insufficiently pure participants in public discourse, they will be more likely to do so. Anyone who pays even casual attention to social media will recognize that we are far from having an established norm along these lines. Current practices are quite the opposite, in fact, as people who are seen as black sheep to some political team are regularly pummeled for their impurities.

There are also norms that could encourage the fruitful expression of changes in view. We could, for example, praise people who express disagreement with much of the rest of their in-group about some issue or event. Or better still, we could adopt more widely the practice of some debating societies of holding in especially high esteem those who can provide clear and compelling explanations about why certain arguments moved them to change their minds. The idea behind these norms is that the independent-minded can serve as helpful models for others to reexamine their own beliefs, even if they don't end up changing their minds along with them.

To this vision of a society of unblocked exits, it might be objected that there is such a thing as too much intellectual independence, and we should be mindful of excess in that direction. We agree that there is something to this worry, but it arises mainly because people confuse being independent-minded with other less admirable traits. To some, the mark of an independent mind is that a person apparently has no firm beliefs, or is always ready to provide a surprising take that bucks all conventional wisdom. If we reward people for these traits, we can expect a profusion of shallow and poorly thought out contributions to public discourse. To be clear, this is not our ideal of a free thinker. Being determined to surprise people or to frustrate expectations is just one more way of letting other people determine the content of your beliefs. A free thinker worthy of the name follows the arguments wherever they lead, and only because they lead there. That is our ideal.

Here's another objection. We have argued that exit blockers err. People shouldn't impose social costs on those who change their minds.

But perhaps we have overstated the case. Communities and associations organized around shared goals, values, and lifestyles are valuable. People find meaning, status, and purpose in these social networks, and presumably because they are sources of value, they should be preserved. One way to hold communities together is to impose some social costs for defection. People *should* feel the sting of turning their backs on the beliefs and values of their communities. Indeed, this may be one resource that communities possess to prevent people from changing their beliefs for bad reasons (such as following some moral fad). Exit-blocking may also be a way to prevent people from changing their minds prematurely—say, after coming across a single smart person arguing for the other side. Furthermore, there may be important accumulated but latent knowledge within the received tradition of a social group, and individuals should be reluctant to turn their backs on their moral and epistemic communities. Social pressure in the form of in-group exit-blocking can be a justified and valuable practice to preserve those communities.

One reply to this objection is simply to deny that such exit-blocking is ever permissible or morally valuable. People should not impose social costs for changing your mind, even for these apparently noble reasons. Only purely individualist epistemic considerations are at issue, and it's morally wrong to sanction people who defect or to try to deter others from defecting. That's one kind of response. A second, and to our minds, more plausible reply to this objection is to note that sometimes social pressure can be a permissible way to preserve a community, but that it matters how and why this pressure is applied and what costs are imposed.

For example, if group members block the exits in order to preen and parade their alleged moral purity, this not only would be morally bad in its own right (as an instance of grandstanding), but also an epistemically unreliable way to preserve the latent knowledge in a moral community, privileging as it does status-seeking over truth-seeking. But if modest, appropriately motivated social costs could be imposed on defectors, there may be some countervailing moral value in preserving social cohesion and a tradition. The challenge, then, would be to articulate the contexts in which such exit blocking is permissible, and what equilibrium exists between the norms against exit-blocking and those in its favor. This is a task we identify but won't pursue here.

#### **4.4 Conclusion**

As people pursue ever more avenues for public discourse about morality and politics, more opportunities arise to have productive conversations. More opportunities also arise to abuse the public square for personal gain. Exit-blockers, we've argued, illegitimately impose social costs on others simply for changing their minds, often because they want to show

off their moral purity, or to feel powerful as they dominate those who publicly admit they've been wrong. This is a bad practice, and its collective effect is to create a less free, more cruel, and epistemically stunted society.

## Notes

- 1 Box 43092, Texas Tech University, Lubbock, TX 79409-3092.
- 2 311 Shatzel Hall, Bowling Green State University, Bowling Green, OH 43403.
- 3 This fear is more common among conservatives, but a majority of liberals and moderates also report sharing it. Notably, the 62 percent figure is up 10 percent from just three years ago. <https://www.cato.org/publications/survey-reports/poll-62-americans-say-they-have-political-views-theyre-afraid-share> [accessed January 21, 2021]
- 4 We don't think that every cost resulting from public speech must be a social cost. Social costs are just a subset of all the costs someone might suffer.
- 5 We say more about the problems grandstanding causes for free expression in (Tosi and Warmke 2021).

## References

- Bicchieri, Cristina. 2016. *Norms in the Wild: How to Diagnose, Measure, and Change Social Norms*. New York: Oxford University Press.
- Cheng, Joey T., Jessica L. Tracy, and Joseph Henrich. 2010. "Pride, Personality, and the Evolutionary Foundations of Human Social Status." *Evolution and Human Behavior* 31 (5): 334–47. <https://doi.org/10.1016/j.evolhumbehav.2010.02.004>
- Joshi, Hrishikesh. 2021. *Why It's OK to Speak Your Mind*. New York: Routledge.
- Marques, José M., Vincent Y. Zyerbyt, and Jacques-Philippe Leyens. 1988. "The 'Black Sheep Effect': Extremity of Judgments towards Ingroup Members as a Function of Group Identification." *European Journal of Social Psychology* 18 (1): 1–16. <https://doi.org/10.1002/ejsp.2420180102>
- Mill, John Stuart. 1989. *On Liberty and Other Writings*. Cambridge: Cambridge University Press.
- Pinto, Isabel R., José M. Marques, John M. Levine, and Dominic Abrams. 2010. "Membership Status and Subjective Group Dynamics: Who Triggers the Black Sheep Effect?" *Journal of Personality and Social Psychology* 99 (1): 107–19.
- Radzik, Linda. 2020. *The Ethics of Social Punishment: The Enforcement of Morality in Everyday Life*. Cambridge, United Kingdom: Cambridge University Press.
- Tosi, Justin, and Brandon Warmke. 2016. "Moral Grandstanding." *Philosophy and Public Affairs* 44 (3): 197–217.
- . 2020. *Grandstanding: The Use and Abuse of Moral Talk*. New York: Oxford University Press.
- . 2021. "Moral Grandstanding as a Threat to Free Expression." *Social Philosophy and Policy* 37 (2): 170–89. <https://doi.org/10.1017/S0265052521000108>