

ARTICLE

The Place of “The Liberty of Thought and Discussion” in *On Liberty*

Dale E. Miller 

I consider whether Mill intends for us to see the arguments that constitute his defense of the “Liberty of Thought and Discussion” in chapter 2 of *On Liberty* as a part of his larger case for the “harm” or “liberty” principle (LP). Several commentators depict this chapter as a digression that interrupts the flow between his introduction of this principle in the first chapter and his exposition and defense of it in the final three. I will argue instead for a reading of *On Liberty* on which chapter 2 is well integrated with the rest of the essay and is in fact where Mill starts constructing his case for LP.

1. Introduction

On Liberty remains widely read, by the standards of philosophical treatises, but its accessibility is only skin deep. It is easy to come away from a first reading of the essay satisfied that one has understood Mill nearly perfectly. Yet on each subsequent reading new puzzles emerge, so that scholars differ sharply over how to read a text that undergraduates find comparatively easy.

My topic here is one aspect of *On Liberty* where such differences appear, namely the relation between Mill’s defense of the “Liberty of Thought and Discussion” in chapter 2 and the remainder of the essay. More specifically, the question that interests me is whether Mill intends for us to see the arguments that constitute this defense as a part of his larger case for the “harm” or “liberty” principle (LP). Several commentators depict this chapter as an awkward *intermezzo*, a digression that interrupts the flow between his introduction of this principle in the first chapter and his exposition and defense of it in the final three. I will argue instead for a reading of *On Liberty* on which chapter 2 is well integrated with the rest of the essay and is in fact where Mill starts constructing his case for LP. The textual evidence in support of this reading is compelling but not conclusive, so the case for it is not open and shut. It is a strong *prima facie* case, though, and I will show that this reading’s leading critics fail to rebut it.

In the next two sections I will prepare the way, by briefly explaining how I understand LP and rehearsing Mill’s arguments for the freedom of discussion in chapter 2 (largely taking for granted a reading of *On Liberty* that I have defended elsewhere (Miller 2010: 113–53)). In section 4, I will present the textual evidence in favor of my reading, which includes key passages that immediately precede and follow chapter 2 and so serve to frame it. Section 5 will consider Jonathan Riley’s contention that Mill sees nearly all expression as “other-regarding” and so outside the scope of the liberty

© The Author(s), 2020. Published by Cambridge University Press

protected by LP. Section 6 will address a different approach to reading Mill, one advanced by first John Skorupski and now Chris Macleod. While not denying that some of the expression that Mill has in mind in chapter 2 might be “self-regarding” and so protected by LP, they still deny that Mill intends for us to see his arguments there as part of his case for that principle.

2. The liberty principle

LP is the keystone of *On Liberty*. In chapter 1, Mill describes asserting “this one simple principle” as the singular “object of this Essay” (1977: 223). While LP may “govern absolutely the dealings of society with the individual in the way of compulsion and control,” Mill’s statement of purpose is not so absolute that we should expect *On Liberty* to do literally nothing other than articulate, defend, and apply it. Nevertheless, if this statement means anything, it is that his treatment of LP constitutes the essay’s core and that everything else is peripheral.

What LP says, at least as a first approximation, is that preventing an individual from directly harming others without their consent is the only consideration that counts in favor of coercively restricting her freedom of action.¹ ‘Harm’, in this context, is a term of art that should be understood as referring to “definite damage, or a definite risk of damage” to someone’s interests (Mill 1977: 282). The idea that Mill conceives of harm in terms of damage to interests originates with John Rees (1960); he points out that because not everything that affects us threatens our interests, this allows Mill to limit the grounds on which social interference can be justified without denying that our lives are intertwined. ‘Interests’, in turn, should be roughly understood to refer to things that are especially valuable or useful regardless of your specific ends in life or sources of subjective happiness, loosely akin to Rawlsian primary goods (Rawls 1971: 62, 92; see also Barry 1965: 174–86). Admittedly, Mill never explicitly defines ‘harm’ in this way (or any other). Yet he connects harm to setbacks of interests, and this definition coheres well with his examples of harms given plausible assumptions about what interests he takes us to have (e.g., Mill 1977: 224–25, 233, 266, 276–77, 279, 285). This is a more restrictive view of how Mill understands harm than, say, Piers Norris Turner’s claim that Mill’s conception of harm is so “expansive” as to encompass any negative consequence (Turner 2014). As Turner reads Mill, the mere fact that some busybody would be displeased by our participating in some activity makes it the case that we harm them by doing so. My view is less restrictive than others, though, such as interpretations that take Mill to equate harm with damage only to a select subset of interests and/or with rights violations (Gray 1996: 57; Brink 2013: 137). The less restrictive Mill’s conception of harm, note, the weaker LP is – the more openings there are that can be used to justify coercive interference.

While Mill never offers a list of our interests, either, he would likely agree with John Locke that they include “life, liberty, health, and indolency of body; and the possession of outward things” (Locke 1955: 17). To this Mill would add security, broadly understood, which in *Utilitarianism* he describes as “the most vital of all interests” (Mill

¹There are various complications here that I will ignore at present. Some concern society’s ability to compel people who are not themselves harming anyone else to participate in protecting others from harm. Others concern issues raised by Ben Saunders (2016), who argues that LP can be more economically expressed as saying that interference is justified only in order to prevent non-consensual harm, whether to the agent or to others. Taking these points into account would not materially alter my conclusions.

1969: 251). He would also add likely privacy, albeit perhaps narrowly construed in terms of the disclosure of personal information; this is implicit, for instance, in his statement that he “would not permit the press to impute, even truly, acts, however discreditable, which are in their nature private” (Mill 1982: 167). And he might well expand the list still further. But we should expect him to be wary of proliferating interests too freely, since the more interests we recognize, the more effects that count as harms and the weaker LP becomes.

More needs to be said about what Mill means by harming someone “directly” than I can say here. However, in general Mill does not believe that the fact that Robert influenced Dahlia to do harm entails that Robert himself harmed anyone, in cases where Dahlia was still able to deliberate and decide for herself. He says, for instance, that my setting a bad example that someone else followed to her own detriment does not amount to my harming her (Mill 1977: 278). Her ruin may be a causal consequence of my bad choices, but not “directly, and in the first instance” (Mill 1977: 225).

Importantly, LP states only a necessary condition for society to be justified in restricting an individual’s freedom of action, not a sufficient one. Mill believes that there are many activities that are sufficiently beneficial that they should be tolerated even though they harm others without their consent, including participating in competitive examinations or trading in a competitive economy (1977: 292–93). The function of LP is simply to restrict the reasons that count in favor of restricting freedom of action to just one, namely the prevention of harm to unwilling others. In the case of any proposed restriction, the benefits that it would yield in terms of harm prevention must be weighed against the costs of restriction.

It can be helpful to think of *On Liberty* as setting out a two-stage decision procedure for deciding whether social coercion is justified – an “interference algorithm.” At points Mill’s own exposition is framed in terms of such a procedure (although the label is not his). In the first step of this procedure, we ask whether the proposed interference would protect anyone except the individual(s) being interfered with from being harmed without their consent. If not, then per LP the interference is forbidden; there is no valid reason in favor of restriction. If so, then we move to the second step, and ask the pragmatic question of whether the costs of the proposed restriction are an acceptable price to pay for the level of harm prevention in question. Only if the answer to this second question is also affirmative should coercion be employed.²

Mill describes conduct that does not directly harm others as “self-regarding” (1977: 281). It lies, he tells us, outside of society’s “jurisdiction” (1977: 276). It is interference with self-regarding behavior that is ruled out at the first step of the interference algorithm. In societies that have reached a certain level of “civilization,” restricting it can never be justified.

Although I will not make a case for this reading of *On Liberty* here, it is hardly idiosyncratic. The proposition that LP disallows everything except the protection of others as a reason for interfering with an individual’s conduct deserves to be called the received view, even if interpreters have construed it differently and a few dissent

²LP does still play a role in the algorithm’s second stage, since it precludes us from counting any considerations other than harm prevention in favor of interference. So the second stage is not an all-inclusive utilitarian calculation. While the two-stage algorithm is a useful expository device, strictly speaking the first stage is dispensable. We might simply say that if an interference would prevent no harm then it cannot be justified because there is nothing to weigh against its costs. See Miller (2010: 129–32).

from it entirely.³ Numerous commentators follow Rees in terms of taking Mill to equate harm with damage to interests, even if they disagree over how precisely Mill understands the notion of interests, and even if Turner and some others dissent.

3. Mill's defense of the freedom of discussion

While the title of chapter 2 refers to the liberty of both thought and discussion, its focus is almost entirely upon the latter. This is unsurprising, given that it is expression and not silent thought that is in danger of suppression. Mill's use of 'discussion' rather than 'expression' is noteworthy. His arguments about the social costs of censorship in chapter 2 only pertain to a proper subset of expression; he does not say or even imply that there are no social costs to silencing other forms of expression, let alone that they should routinely be silenced, but in detailing the costs of censorship he restricts his attention to this particular sort. What is unclear is precisely how to delineate this subset. Skorupski glosses discussion as "dialogue," which he characterizes as "unconstrained discourse between rational people" whose "internal goal is right action and right belief – truth in a wider sense in which truth is what practical as well as theoretical reason aim at" (1989: 371). In a similar spirit, Turner writes that "'Discussion' is Mill's consistently employed word for joint, reasoned engagement on some (usually public) matter, governed by norms of truth, fair play, and sincere attention to the general good" (Turner 2013: 105). Macleod, more broadly, takes Millian discussion to encompass all assertions of opinions that are "truth-apt." This includes moral and political opinions, but still excludes "jokes," "poetry or literature, painting or music," and "acts of protest or support" (Macleod 2021: 14).⁴ I will remain agnostic on how precisely Mill delimits discussion, noting how these differences between definitions could affect my account when relevant.⁵ However, it is worth bearing in mind a criticism that Riley advances of Skorupski's reading on this point, namely that the more norm-laden and rarified we take Mill's conception of discussion to be, the narrower and therefore weaker the arguments of chapter 2 become (Riley 2005: 152). Riley's worry may be even more pressing when directed against Turner, according to whom Millian discussion requires commitments not only to truth and discursive fairness but also to the general good. (Does debate no longer qualify as discussion if participants aim to discover the truth at any cost?)

Mill's case for the freedom of discussion is frequently described in terms of the "marketplace of ideas," although this term would not be coined until the twentieth century.⁶ As a first approximation, we can say that this case is epistemic: the freedom of discussion is a necessary means for the effective discovery and dissemination of truth. Yet this account, while accurate as far as it goes, is incomplete.

³One dissenter is Daniel Jacobson, whose novel account of Mill's theory of the freedom of expression rests on a revisionary view of the nature of LP and its role in Mill's liberalism (Jacobson 2000, 2017). I will not set out my differences with Jacobson's reading of chapter 2 here because doing so would require spending too much time litigating our deeper disagreements about LP.

⁴In his contribution to this symposium Macleod expands the notion of discussion to include expression that helps us think through truth-apt propositions even if it is not truth-apt itself.

⁵At the end of chapter 2, Mill touches upon some norms governing discussion and the perils of trying to enforce these norms. While he acknowledges that the "true morality of discussion" requires candor and fairness to our opponents, it is not clear that he means to imply that expression that fails to reflect this morality thereby falls outside the scope of discussion (1977, 258–59).

⁶While I will use this metaphor here, it has been criticized both in general and as a characterization of Mill's reasoning (e.g., Gordon 1997, Nordenstreng 2016).

Suppose that society wants to suppress some doctrine. Mill says that there are three possibilities. First, the doctrine might be true, in which case the epistemic cost of suppression is clear (Mill 1977: 229–43). Second, it might be partly true, and Mill maintains that it is only through free discussion that the kernels of truth in different views can be separated from the chaff (1977: 252–57). Third, it might be false. In this case, Mill contends, we are only positioned to be confident that a view is false if those who hold it are free to say whatever they can in its defense, so that we can hear and assess their arguments (1977: 229–31). In addition, he famously emphasizes that when people know that they do not need to worry about their beliefs being challenged then they tend to hold them as “dead dogma” (1977: 243–52). They neither concern themselves with the grounds for these beliefs nor live in ways that reflect a deep commitment to their truth. So for society to silence those who would advocate false views negatively affects how people hold true ones. What a purely “cognitive” or epistemic reading of the chapter leaves out is that Mill believes that discussion matters not only for the discovery and dissemination of truths but also for how we incorporate those truths that we have discovered into our lives, including how they shape our feelings and motivations.

Mill’s arguments in chapter 2 show that insofar as it is valuable for people to have true beliefs and (where applicable) to base their lives on these beliefs, we have reason to value the freedom of discussion. We must recognize, though, that the arguments of chapter 2 only count in favor of freedom of discussion *pro tanto*. They leave open the possibility that countervailing considerations will sometimes justify restriction.

4. The continuity interpretation

My proposal for how we should understand the contribution of these arguments to Mill’s overall argumentative strategy in *On Liberty* is straightforward. It is that in chapter 2 Mill argues that one specific variety of self-regarding conduct should also be left free from social interference or censorship, namely self-regarding “discussion.” He then argues, in chapters 3–5, that the same is true of self-regarding conduct generally. So there is a straightforward progression of thought throughout the essay which chapter 2 advances rather than interrupts. I will call this the “continuity” interpretation.

While he admittedly could have been even more explicit, Mill is remarkably clear about how chapter 2 fits within the essay as a whole in the concluding paragraph of chapter 1. This paragraph reads in part:

- (1) *It will be convenient for the argument, if, instead of at once entering upon the general thesis, we confine ourselves in the first instance to a single branch of it, on which the principle here stated is, if not fully, yet to a certain point, recognised by the current opinions. This one branch is the Liberty of Thought: from which it is impossible to separate the cognate liberty of speaking and of writing. . . . Those grounds, when rightly understood, are of much wider application than to only one division of the subject, and a thorough consideration of this part of the question will be found the best introduction to the remainder.* (1977: 227 [emphasis added])

(I have numbered this extract for future reference.)

Mill says in the italicized sentences that the liberties of thought and discussion are one area in which the principle that he has just stated – LP – already enjoys some

measure of social acceptance. This strongly implies that he sees these liberties being grounded on LP. Moreover, “the argument” to which the first sentence refers is the essay’s overarching case for LP. So Mill frames the specific arguments in chapter 2 as a component of this case, a part of a larger whole. The continuity interpretation simply takes Mill at his word about all of this. While it is helpful that Mill spells this out, it is in any case what we should expect. It would be distinctly odd if, immediately after declaring that his single official ambition in *On Liberty* is to win support for LP, he launched into an extended discussion that brought him no closer to this end. Indeed, it would be rather surprising for him to devote an entire chapter – one nearly double the length of any of the rest, regardless of its placement – to ruminations that did nothing to help him realize that ambition.

Mill’s arguments for why the freedom to participate in discussion should rarely if ever be restricted do not have LP as a premise, but they confer support on LP by showing that its implications in this important area of conduct are correct. It reaches the right results. This manner of arguing depends upon two assumptions. The first is that as a general rule – in the great majority of cases, if not literally all – discussion is self-regarding. The second is that any exceptional cases in which discussants should be silenced are also exceptions to this general rule. If Mill’s arguments for the freedom of discussion in chapter 2 are intended to help to establish LP, then he must accept these claims.

With respect to the first of these points, Mill never explicitly says that he believes that discussion is self-regarding in at least most instances. Yet (1) is a compelling contextual clue that he does. And this belief is a plausible one, from his perspective, at least so long as I am correct that he has a relatively restricted conception of harm. Many of the negative effects on people to which discussion can give rise, for example, offense or hurt feelings, do not involve damage or risk of damage to interests. (This belief would obviously be wildly implausible from the standpoint of someone who conceives of harm in the expansive manner that Turner ascribes to Mill.) I should note here that I am presupposing that the discussion in question does not run afoul of ordinary time, place, and manner restrictions. However precisely we characterize discussion, your participating in it can certainly be harmful if you are wielding a megaphone in the wee hours just outside my bedroom window.

With respect to the second point, there may be some disagreement about whether Mill believes that discussion can ever be restricted; the differences between interpreters over his definition of ‘discussion’ come into play here. He does give a few illustrations of when expression can justifiably be restricted, the best known of which appears early in chapter 3:

[O]pinions lose their immunity when the circumstances in which they are expressed are such as to constitute their expression a positive instigation to some mischievous act. An opinion that corn dealers are starvers of the poor, or that private property is robbery, ought to be unmolested when simply circulated through the press, but may justly incur punishment when delivered orally to an excited mob assembled before the house of a corn dealer, or when handed about among the same mob in the form of a placard. (1977: 260)

Here we are evidently to imagine a crowd of people so agitated that they will not be able to think clearly about whether to react violently if they hear the opinion in question. Since the ensuing riot will not be mediated by their own deliberation, the harm

caused by the riot can be ascribed to the speaker's act. Do the circumstances of this expression of opinion take it out of the realm of discussion? Apparently not, as Macleod takes Mill to define 'discussion', insofar as it is still a truth-apt moral *cum* political proposition; quite possibly so, as Skorupski and Turner take him to define it, since in context it does not appeal to its hearers' rational faculties and its utterance does not promote the general good. For my purposes, happily, I do not believe that anything turns on this. If it is discussion, then it is an exception to the general rule that discussion is self-regarding.⁷ Indeed, by introducing the corn-dealer scenario immediately after his defense of the liberty of discussion Mill appears to be drawing a contrast between typical cases in which expression is self-regarding and exceptional cases in unusual circumstances in which it is not, and this rhetorical move offers further support to the continuity reading.

In terms of the interference algorithm, the continuity reading of chapter 2 claims that Mill sees most censorship of discussion as being ruled out in the first step. If any discussion is harmful, sometimes the second step will be reached. Any such cases would be comparatively unusual and marginal, however – enough so that they can be ignored in chapter 2 itself.

5. Riley's "laissez-faire with exceptions" interpretation

Jonathan Riley offers what appears at first glance to be an especially nice statement of the continuity interpretation in the original edition of an introductory text on *On Liberty*:

All thought is self-regarding, [Mill] makes clear, and virtually all expression is reasonably classified as self-regarding as well. In special circumstances, however, the expression of an opinion cannot be treated as self-regarding because it has 'at least a probable connexion' to an act which is seriously harmful to others. In those special cases, where expression is reasonably taken out of the self-regarding sphere, there is no moral right to liberty. . . . It is important to keep in mind that Mill excludes some expression from the self-regarding sanctuary, so that his liberty principle does not pretend to grant absolute protection to all expression. With that caveat, his prescription of absolute liberty for what may be termed 'self-regarding expression' is compatible with his claim that society has legitimate authority to control expression in the special cases. (Riley 1998: 56)

In fact, however, out of context this passage gives a very misleading impression of Riley's interpretation. When he says that most expression is "reasonably classified as self-regarding," what he actually means is that Mill believes that expression should be treated "in 'almost' all instances *as if* it were self-regarding" – that is, it should

⁷Other examples that Mill gives of expression's being justifiably restricted are similar in that, while there may be some question as to whether the expression in question qualifies as discussion, there is no doubt that he considers it harmful. These examples include a somewhat puzzling footnote in which Mill allows that someone can justly be punished for instigating "Tyrannicide" in a "specific case . . . but only if an overt act has followed, and at least a probable connexion can be established between the act and the instigation" (1977: 228). The puzzles here include, first, what Mill would count as instigation in cases (unlike the corn-dealer scenario) in which the audience is able to deliberate for themselves about whether to act on the speaker's or author's suggestion and, second, why he rules out the possibility of prior restraint while still allowing punishment after the fact.

be left unrestricted – despite the fact that it “can harm others” and so “[s]trictly speaking . . . is legitimately subject to social control” (Riley 1998: 49–50). Expression’s presence in the “self-regarding sanctuary” is merely honorary. Riley summarizes his view in a revised version of this text by saying that in many cases Mill supposes that we can “reasonably pretend” that expression is self-regarding (Riley 2015a: 75). So we must consider two questions: Why does Riley conclude that Mill sees all or almost all expression as potentially harmful? And given that he reaches this conclusion, why does he further hold that Mill believes that we should still be very reluctant to engage in censorship?

(1), taken together with Mill’s more general statement about the centrality of LP to *On Liberty*, is the chief evidence in favor of the continuity interpretation. (1), though, is admittedly somewhat ambiguous: is it only liberty of thought that Mill sees as “one branch” of the wider liberty entailed by LP or liberty of discussion as well? Another passage might appear to answer that question:

- (2) The liberty of expressing and publishing opinions may seem to fall under a different principle, since it belongs to that part of the conduct of an individual which concerns other people; but, being almost of as much importance as the liberty of thought itself, and resting in great part on the same reasons, is practically inseparable from it. (1977: 225–26)

It is on the strength of (2) that Riley rejects the continuity interpretation. He takes Mill to be telling us explicitly in this passage that discussion, and expression more generally, is never self-regarding; he regards Mill’s use of the phrase “concerns other people” to be a term of art that signifies not merely that expression aims at communicating with other people but that it has the potential to harm them (Riley 2005: 147). This is true, Riley maintains, even of expression that constitutes paradigmatic instances of discussion, since “Just as sellers can harm others by driving them out of business in an open competitive market, for example, speakers can harm other speakers by discrediting their opinions and perhaps injuring their reputations in a fair and open debate” (2005: 162; see also 149). Riley, note, takes Mill to identify harm with “perceptible damage” to others, which is evidently a more restrictive understanding than Turner’s but less so than mine (2005: 156).

Rather than seeing expression as generally self-regarding, Riley proposes instead that Mill sees it as closely analogous to participation in market activity. Expression is more or less entirely within society’s jurisdiction, Riley takes Mill to believe, just like buying and selling. Like regulations on trade, proposed regulations on expression always survive the first step of the interference algorithm, and if they are ruled out it is only on the pragmatic grounds that come into play at the second step. Enough regulation of expression is ruled out on this basis that, as with trade, society’s approach should be largely *laissez-faire*; “Riley’s Mill” is no more a proponent of extensive censorship than of extensive economic regulation. Indeed, Riley contends, expression “is virtually never the object of expedient coercion” (1998: 220). But rather than our having an “unqualified” moral right to speak and publish, the sort of right that LP generates and that society is obligated to honor, we have only as much of a right as society chooses to give us – a “qualified” right (Riley 2005: 162).

Two objections to Riley’s interpretation may immediately suggest themselves. The first is that it may seem to make the organizational structure of *On Liberty* a muddle by depicting chapter 2 as a digression from Mill’s stated singular objective of arguing

that the self-regarding sphere should be free from interference. The second is that Riley ignores a key phrase in (2). In contrast with his definiteness about free trade, Mill says that expression “may seem” to fall under another principle; the implication is that we should not take this appearance at face value. Indeed, we do not even need to get to the next full stop before we encounter the expected “but,” after which Mill indicates that thought and discussion are more closely akin than they might initially appear. Remarkably, Riley never grapples with nor even acknowledges that “may seem”; he consistently restates (2) so as to suggest that Mill baldly declares that the expression of opinions *does* fall under a principle other than LP (Riley 2005: 147, 149, 176).

However, Riley makes a point that goes some way toward addressing these objections. He directs our attention to Mill’s assertion that restrictions on the sale of certain items can violate LP, the fact that trade is a social act notwithstanding. This is true when the restrictions cannot be justified at the second stage of interference algorithm and the only plausible explanation for them is to prevent prospective buyers from using them for self-regarding purposes. In such cases, Mill says, “These interferences are objectionable, not as infringements on the liberty of the producer or seller, but on that of the buyer” (1977: 293; see also 288).⁸

In a like way, Riley claims, Mill believes that when restrictions on the freedom of expression cannot be justified in terms of harm prevention and can only plausibly be explained as attempts to prevent certain views from being read or heard, they violate LP. More specifically, they abridge the liberty of those who would like to read or hear these views, who are being denied access to the materials for self-regarding thought:

[T]he central principle implies that the individual should have complete liberty to hear another’s opinion if the opinion can be heard without harm to third parties. Society has no legitimate authority to ban the expression of such an opinion, although it may expediently regulate the time, place and manner of expression . . . (Riley 2005: 163)

This gives Riley an easy way to account for Mill’s description of the liberty of discussion as “practically inseparable” from that of thought. Riley may take this claim to apply only to a proper subset of expression, namely the expression of opinion. However, he construes the notion of opinion quite broadly, broadly enough to encompass indecent novels and pornographic films (Riley 2005: 164).⁹

⁸I take Mill’s view to be that the rights of buyers play no role in determining when restrictions on the sale of items are unjustified. It is only when we have already determined that restrictions cannot be justified in terms of harm prevention, via the interference algorithm, that we can conclude that they represent paternalistic or moralistic attempts to prevent access to these items altogether. Riley, in contrast, seems to take him to hold that buyers’ rights preclude even highly expedient restrictions if they would make it impossible to buy commodities with any self-regarding uses (2005: 164). Since machine guns and even grenades can be used for self-regarding entertainment, as numerous online videos attest, as Riley reads him Mill might permit the sale of all arms short of weapons of mass destruction (Riley 2005: 165).

⁹The apparent disagreement between Riley and myself described in the preceding note carries over here. I take Mill’s reasoning to imply that when the expression of opinion *is* harmful, the question of whether restrictions are warranted in terms of the interference algorithm is prior to that of whether these restrictions violate the rights of prospective hearers. Riley appears to think otherwise, taking these rights to require that there be some time, place, and manner in which any opinion can be expressed as long as it can be

Riley does not advance the claim that unwarranted restrictions on the expression of opinion violate the rights of prospective audience members as a way of addressing the initial objections to his reading that I raised, but he can put it to this use. Doing so lets him say that even though discussion and expression more generally are not self-regarding activities, the subject matter of chapter 2 is not completely divorced from LP. When restrictions on discussion cannot be justified in terms of the interference algorithm, and so are exposed as attempts to prevent certain opinions from being aired altogether, someone who wants to hear those opinions will be entitled to complain that a right of hers that is grounded on LP is being violated. Chapter 2's arguments for the freedom of discussion do not contribute to establishing LP, on Riley's reading; they are analogous to economic arguments for why regulations on trade are not cost-effective. So as Riley reconstructs *On Liberty* the essay does not have the structure that we might have expected, one where the arguments in the lengthy second chapter constitute part of the case for LP. Still, it is at least more cohesive than it might have initially appeared. Invoking this claim also gives Riley a way of explicating the troubling "may seem." Mill could be saying that while the expression and publication of opinion may appear to fall under a different principle from LP, and in their own right they do, on closer inspection it turns out that there is an unqualified right grounded on LP in play. It just belongs to prospective audience members, not presenters.

These possible replies notwithstanding, however, strong objections to Riley's interpretation remain. First, Riley's reading is difficult to reconcile with the passage in chapter 4 where Mill introduces the "natural penalties" that a person may permissibly "suffer . . . at the hands of others for faults which directly concern only himself":

We have a right . . . to act upon our unfavourable opinion of any one, not to the oppression of his individuality, but in the exercise of ours. . . . We have a right, and it may be our duty, to caution others against him, if we think his example or conversation likely to have a pernicious effect . . . (Mill 1977: 278)

Here Mill characterizes a person's "conversation" as directly concerning only himself despite its probable "pernicious effect." He cannot intend to suggest that this is true of every possible conversation – not one involving slanderous lies, for instance. He has something more specific in mind: we are to imagine a person who is a spendthrift or is otherwise imprudent, who by example or conversation might influence someone else to develop the same trait. Mill counts this conversation as self-regarding because its deleterious effect, like that of behavior that sets a bad example, is not direct. While we cannot punish this individual for his conversation, we do have a right to advise our friends to avoid him. This is an "unqualified" right. In context, Mill is explaining that our right to respond to people's self-regarding behavior with the natural penalties is itself entailed by LP; he describes the employment of the natural penalties as an exercise of our own "individuality." One sentence therefore contains two examples of expression – the original conversation and the warning – that Mill treats as self-regarding. Riley's interpretation cannot easily account for these examples. In fact, Riley explicitly denies that Mill's theory gives one person a right to "circulate her sincere and informed opinions of another's character and activities" when this might "seriously damage that person's reputation" (Riley 2005: 156). But this is precisely what warning

"consumed" by an audience "without direct and immediate harm to others," even if the interference algorithm calls for preventing all possible expressions (Riley 2005: 149).

others about someone's self-regarding character flaws and pernicious conversation amounts to, which Mill says in no uncertain terms that we can have a right to do.

Second, and more significantly, for Riley's reading to be plausible he must have a satisfactory account of how expression generally, and discussion in particular, is always or nearly always harmful to others. Yet the account that Riley offers is both problematic in its own right and un-Millian in spirit. He gives a hostage to illiberal fortune in construing Mill's conception of harm so broadly as to include discrediting others' opinions. We can discredit the opinions of others not only through our words but through our actions. The view that Riley ascribes to Mill therefore has the following surprising implication: should anyone assert that it is impossible to ϕ , then even if openly ϕ ing would previously have been self-regarding this is no longer the case. To publicly ϕ would be to discredit the opinion and perhaps damage the reputation of whoever made this assertion, so one person's opining that ϕ ing is impossible creates a reason for society to prevent anyone else from ϕ ing (except perhaps in secret), even when no such reason existed previously. If any homophobe has ever claimed that gay couples cannot form lasting romantic relationships in the same way that heterosexual couples can, for instance, then Riley's reading entails that she is harmed by the choice of a gay couple to remain in an openly loving long-term relationship. In fact, merely by absurdly and even insincerely declaring this to be impossible, she can bootstrap into being a reason for society to interfere. This illustrates how LP can be dramatically weakened by an expansive conception of harm, a point that Riley has vigorously made himself in criticism of Turner (Riley 2015b).

As an alternative to Riley's reading of (2), the proponent of the continuity interpretation can say that Mill is there acknowledging that the fact that discussion concerns others in the sense of being aimed at communicating with them and having the potential for influencing them might at first glance seem to mean that it also concerns them in the sense of threatening harm. But, he concludes, on closer inspection it turns out not to, at least in the great majority of cases. Even when its influence is not entirely anodyne, expression's effects are not usually harmful in the sense relevant to LP.

While this *is* what I want to say is going on in (2), I admit that Mill could have made this clearer and that it would be better for my reading if he had. Yet on balance Riley's interpretation fits less well with the text on the whole than the continuity interpretation and ascribes a less attractive and less liberal view to Mill. But I do not want to overstate the extent to which Riley and I disagree. We remain in rough agreement, I think, about when Mill would want society to interfere with someone's expression. We only disagree about the grounds on which he would make this determination. Importantly, someone who subscribes to the continuity interpretation can agree with Riley that Mill sees some expression (and perhaps even some discussion, depending on how this is defined) as harmful – and not only in cases as explosive as the corn-dealer scenario. They can further agree that Mill would believe that some of this harmful expression is worth tolerating.¹⁰

¹⁰Riley also offers very effective criticism of K. C. O'Rourke's account of Mill's theory of the freedom of expression (O'Rourke 2001; see Riley 2005: 152–59). O'Rourke's interpretation superficially resembles mine, since he argues that Mill takes LP to entail a right to freedom of expression, but we understand LP quite differently. O'Rourke argues that while LP is only a very weak anti-paternalism principle, it still somehow entails a nearly unqualified right to express opinions (O'Rourke 2001: 120, 122). Riley shows this to be incoherent.

6. Skorupski, Macleod, and the “Principle of the Liberty of Expression”

In contrast with Riley, Skorupski and Macleod are willing to agree that many proposed restrictions on discussion may be ruled out by LP at the first stage of the interference algorithm. Nevertheless, like Riley they take Mill’s arguments in chapter 2 to be concerned with what goes on in the second stage, so their approach is also at odds with the continuity interpretation.

While Skorupski and Macleod write separately and may have some minor disagreements, the similarities between their accounts of Mill’s theory of the freedom of expression make it appropriate to describe them as sharing one interpretation. Their primary aim is to explicate the arguments that Mill gives in chapter 2, and they write about this at considerably greater length and depth than my earlier cursory remarks. When it comes to thinking about chapter 2’s place in *On Liberty*, they largely discount (although they do not entirely ignore) (1) and the other passages outside of chapter 2 that I have emphasized. In contrast with Riley, they also put no special weight on (2). Instead, they ask what role LP plays in the arguments in chapter 2 and how these arguments compare to arguments in other parts of the text.

Skorupski and Macleod see chapter 2 as Mill’s attempt to defend a principle other than LP, which Skorupski calls the “Principle of Liberty of Expression” (PLE) and Macleod the “Freedom of Discussion Principle” (FDP) (Skorupski 1989: 369; Macleod 2021: 3). PLE specifically “prohibits restrictions on honest dialogue” (Skorupski 1989: 373). Or, perhaps more carefully, it rules out the prohibition of dialogue based on consequences that are “routed through a recipient’s rational response” as long as the speaker does not aim to be misleading and has duly considered what she says (Skorupski 1989: 370–72). Macleod takes FDP to say that “there should be no interference with the discussion of any opinion” (2021: 8).

Both Skorupski and Macleod acknowledge that some of the honest dialogue or discussion that falls within the scope of PLE/FDP will be self-regarding, but they maintain as well that some substantial portion will not. This is important, because it is only when the expression in question is harmful that PLE/FDP comes into play. When expression is self-regarding, LP already protects it from interference. When honest dialogue or discussion falls within society’s jurisdiction, however, then PLE/FDP figures in the second stage of the interference algorithm. Macleod puts the point this way:

Where the assertion of a proposition is also an action which causes harm – as we might think of assertions which amount to invasions of privacy, incitement, and so on – interference, therefore, is an option. Nonetheless, the Freedom of Discussion Principle retains force, providing independent considerations that count, albeit non-decisively, against such interference. The Freedom of Discussion Principle, to put this another way, provides extra reasons which must be taken into account in calculating whether interference is warranted *all things considered* – and as such the bar for interference in such cases is raised. (2021: 19)

Skorupski likewise says that PLE “spells out further safeguards,” safeguards above and beyond those offered by LP, “defining limits within which a certain type of discourse – discussion, honest dialogue – is protected from social sanction” (1989: 373–74).

A proponent of the continuity interpretation can go a certain distance with Skorupski and Macleod. She can agree, as I have noted, that sometimes discussion

may not be self-regarding. In those cases, decisions about whether to permit or stifle it will need to be made at the second stage of the interference algorithm, and the epistemic considerations that Mill raises in chapter 2 must figure in the calculation as counting in favor of permission. The utility of expressing this by positing the existence of an additional principle to which Mill himself never refers may be debatable, but the continuity interpretation does not rule this out.

However, Skorupski and Macleod further contend that Mill does not take himself to be advancing his case for LP in chapter 2. His arguments there are meant to bolster PLE/FDP and nothing more. Chapter 2 “has a spliced-in appearance,” Skorupski writes, “as though Mill had reflected on liberty of expression at a remove from the framework of the Liberty Principle as such” (1989: 369). This contention obviously is anathema to my reading. It reflects, I believe, mistaken expectations for what sort of arguments Mill would make in chapter 2 if the continuity reading were correct.

Macleod, for instance, observes that:

nowhere in his defense of the Freedom of Discussion Principle does Mill even *mention* the Harm Principle – the notion of harm plays no role whatsoever in the argument of Chapter Two. As the materials were certainly at hand, we might have expected Mill to offer the following derivation in Chapter Two:

[1, Harm Principle]: the only reason we are warranted in interfering in an action is to prevent harm to others.

[2, No Harm in Discussion Claim]: discussion never causes harm,

Therefore: there should be no interference with discussion. [Freedom of Discussion Principle]

Mill pointedly does *not* make this argument, however. (2021: 17–18)

I agree with Macleod that Mill does not make this argument, but I believe he misdiagnoses *why* Mill does not make it. He writes that it is “impossible to resist the conclusion that Mill rejected the argument – either because he thought it unsound on the grounds of the No Harm in Discussion Claim, or because he thought it missed something important.” Macleod’s explanation seems to be that the argument misses the independence of FDP from LP: FDP is so far from being a consequence of LP that it plays an important role in protecting discussion only in those instances in which LP does not protect it, and it rests on distinct grounds.

In contrast, I submit that the reason that Mill does not give this argument is not that he sees it as unsound but rather that for him to give it would be for him to treat LP dogmatically.¹¹ Recall, he only introduces LP near the end of chapter 1; at this point in the inquiry, he has given his audience no reason to accept it. To make it the major premise in a syllogism at this point would beg the question. Showing that we are entitled to deploy LP as a premise is the *raison d’être* of *On Liberty*; what he must do in the essay itself is to attempt to persuade readers that LP’s implications

¹¹Or at least this is the main reason. Mill *would* reject the No Harm in Discussion Claim as being somewhat too strong. However, the addition of a qualifier like “in most cases” would weaken this premise to the point where he could accept it, and the same qualifier might be added to the conclusion. Yet Mill still could not make the argument at this stage of his dialectic even with these qualifications.

are ones they can accept by offering reasons why its social recognition would tend to advance the permanent interests of man as a progressive being.

And so he does. Again, the continuity reading depicts him as starting this process by showing that it is desirable for one field of conduct that is generally self-regarding to remain largely unimpeded, with the only exceptions occurring in (some of) the comparatively unusual cases where it is *not* self-regarding. I certainly agree with Macleod that Mill's arguments in chapter 2 also have some force in cases in which the discussion that might be restricted would be harmful. Discussion's contributions to the discovery of truth and its integration into our lives must be factored into the weighing of reasons in the second step of Mill's algorithm. If Macleod finds it helpful to invoke an additional *pro tanto* principle to make this point, I will not object.

But if we are going to say that Mill is arguing for FDP in chapter 2, then we must also say that he regards establishing FDP as a major contribution to the case for LP. The considerations that ground the former are among those that ground the latter. FDP, we might say, serves as evidence for LP – or, perhaps better, it summarizes an important subset of the evidence for it (even if it also does more than this). I have already acknowledged that Mill could have been even more transparent about his argumentative strategy. This would have involved making some mention of harm in chapter 2, by making explicit that in general discussion does not threaten anyone's interests and explaining why this is so. But he would have done this in addition to making the arguments he gives, to make their connection to LP more obvious; he would not have been able to replace those arguments with a simple deduction. He argues *for* LP, not *from* it. As he must.

According to Skorupski, in turn, it is the fact that Mill regards discussion as a matter of collective epistemic benefit rather than individual right that divides chapter 2 from what comes after (Skorupski 1989, 375). The “heart” of PLE is not “the individual right of the expression's author,” but “the *social* importance of dialogue.”

Liberty of expression is not a special case of the Liberty Principle, nor does it mainly flow from the same source. [W]hereas the Liberty Principle is founded on the importance of giving individuality its legitimate scope, the deepest justification for the Principle of Liberty of Expression is that it gives a hearing to the communal voice – that to which we respond in common . . . [Mill] is not presenting a special case of the Liberty Principle; he is defending the dialogue model by appeal to its internal goal. (Skorupski 1989: 376)

So Skorupski apparently attributes the “spliced in” character of chapter 2 to two factors, which he seems to see as related. One is that Mill's way of arguing in chapter 2 is different from that in the other chapters, in not being concerned with the benefits to the individual of her own possession of liberty. The other is that it does not treat the liberty in question as a matter of individual right.

Skorupski is not wrong to observe that, despite the last sentence of (1), Mill's arguments in chapter 2 do not entirely parallel those in the later chapters. In chapter 2, he focuses chiefly on the fact that permitting free discussion is the best way for people generally to arrive at the truth, which is common to all. In the later chapters, in contrast, the emphasis is on the fact that each individual's enjoyment of freedom of action more generally is necessary in order for her own individual differences to be accommodated.

We should not, though, overstate the distance between these chapters. A version of the marketplace of ideas recurs in chapter 3, in Mill's suggestion that one reason for

allowing “experiments in living” is that some people may turn up new ways of living that will be suitable for widespread if not universal adoption. The masses may “learn something” if they allow the most original members of the community to develop their individuality, Mill says; in this way, “better modes of action, and customs more worthy of general adoption, may be struck out” (1977: 267, 270). Moreover, Mill could have introduced more of the considerations that he raises in chapter 3 earlier. He might, for instance, have considered how engaging in self-expression contributes to the development of individuality.¹² This would have allowed him both to say more about the costs that suppression imposes on the particular individual who cannot freely express herself and to defend the freedom to engage in forms of expression beyond discussion, which might have made for a better chapter. However, while Mill could have broadened his approach in chapter 2, this would not have altered the chapter’s structural role in the essay as a whole.¹³

Furthermore, Skorupski’s denial that at its “heart” chapter 2 is concerned with an individual right to freedom of expression on the part of speakers and authors seems to be mistaken. Mill is arguing for the existence of this right; showing that its recognition yields a collective social benefit is his means of doing so. For the utilitarian Mill, *any* argument for an individual right is ultimately going to be grounded on a collective benefit in terms of aggregate happiness. In the case of his argument for the right to engage in discussion, more simply rests on an intermediate collective good.

Skorupski and Macleod have done important work in helping us understand exactly how Mill’s arguments in chapter 2 function. In thinking about this chapter’s relationship to the rest of the essay, however, they may focus too narrowly on those arguments and pay too little attention to how Mill frames them in (1) and other passages. It is likely true that if one were to read chapter 2 in isolation, one would come away with the impression that Mill was arguing for a principle like PLE/FDP and nothing more. But Mill did not write the chapter in isolation. A plausible reading of the whole of *On Liberty* is available on which it does not have the disjointed structure that Skorupski and Macleod portray. They have not offered any compelling reasons for us to forgo it.

7. Conclusion

So, to summarize: Several passages that fall outside of chapter 2 – most notably the preview of that chapter that Mill gives at the end of chapter 1, but also his descriptions of *On Liberty* as being organized around the defense of a single principle and presentation of the corn-dealer scenario at the start of chapter 3 – lend considerable support to what I have called the continuity interpretation. Mill’s failure to spell out his argumentative approach in detail and state clearly that he considers most discussion to be self-regarding means that some uncertainty about whether this interpretation is the right one is ineliminable. Nevertheless, we have more reason to accept it than either of (what I take to be) the most plausible alternative accounts of what Mill is doing in chapter 2.

Of course, just because the continuity interpretation of chapter 2 seems to be the correct exegesis does not mean that the view that it ascribes to Mill is correct. My

¹²O’Rourke does emphasize that Mill considers being exposed to the expression of a wide variety of *others* to be a precondition of the cultivation of individuality (O’Rourke 2001: 79).

¹³Macleod discusses the support that chapter 3 could give to the freedom of expression generally in his contribution to this symposium.

focus in the foregoing has been on how to interpret Mill; I have not asked whether most discussion (or, more generally, most expression) really *is* self-regarding. This supposition has been challenged by, *inter alia*, those who claim that what we might loosely refer to as “hate speech” can be harmful and can even constitute “linguistic violence” or “assaultive speech” (Matsuda et al. 1993: 1).

In a recent blog post, Martin Lenz (2019) offers a balanced and accessible account of a distinction between two views of language, one on which we use language to entertain thoughts in a way that is detached from action and one on which expressing thoughts in language *is* action. On the latter view, expression – like any other species of action – can be harmful. As Lenz rightly says, there is truth in both views; we cannot sensibly embrace one to the complete exclusion of the other. As I read Mill, however, he is fairly close to the “detached” end of the continuum – closer, at least, than Lenz. While he would not deny that to verbalize or to publish thoughts is to act, and that these actions are morally evaluable, in the great majority of cases he thinks that there are categorical differences between these actions and others in terms of their ability to cause direct harm. We know from the corn-dealer example that he acknowledges that expression *can* sometimes be harmful “in the first instance,” but that case features a very specific set of circumstances, and one in which expression is harmful as a result of its intimate connection to physical violence rather than entirely in its own right. Even if he does not believe that expression can only be directly harmful where just this type of connection exists, his choice of the example strongly suggests that he sees cases in which expression is directly harmful as infrequent departures from the norm.

Mill may greatly overestimate how much discussion qualifies as self-regarding, either because he fails to appreciate just how often discussion can threaten the interests of others or because this interest-based conception of harm is too narrow.¹⁴ If so, then his arguments in chapter 2 may do less to advance his case for LP than he believes. But the balance of evidence strongly favors a reading of Mill on which *he* takes those arguments to constitute an important part of that case.¹⁵

References

- Barry, Brian. 1965. *Political Argument* (Berkeley: University of California Press).
- Brink, David. 2013. *Mill's Progressive Principles* (Oxford: Oxford University Press).
- Gordon, Jill. 1997. John Stuart Mill and the ‘Marketplace of Ideas,’ *Social Theory and Practice*, 23: 235–49.
- Gray, John. 1996. *Mill on Liberty: A Defence*, 2nd ed. (New York: Routledge).
- Jacobson, Daniel. 2000. Mill on Liberty, Speech, and Free Society, *Philosophy and Public Affairs*, 29: 276–309.
- Jacobson, Daniel. 2017. Mill on Freedom of Speech, in *A Companion to Mill*, ed. by Chris Macleod and Dale E. Miller (Malden: Wiley), pp. 440–53.
- Lenz, Martin. 2019. Words as Weapons? Free Speech Requires Responsible Speech, *Handling Ideas* (23 April), <<https://handlingideas.blog/2019/04/23/words-as-weapons-free-speech-requires-responsible-speech/>>.
- Locke, John. 1955. *A Letter Concerning Toleration*, 2nd ed. (Indianapolis: Bobbs-Merrill).
- Macleod, Chris. 2021. Mill on the Liberty of Thought and Discussion, in *The Oxford Handbook on Freedom of Speech*, ed. by Adrienne Stone and Frederick Schauer (Oxford: Oxford University Press), pp. 3–19.

¹⁴Melina Bell’s contribution to this symposium cites evidence concerning the effects of certain kinds of insults on their subjects’ health that lends support to attributing the first of these failures to Mill.

¹⁵For valuable comments I am grateful to an audience at the 2019 Philosophy, Politics, and Economics Society Meeting, Leonard Kahn and his students at the Loyola University of New Orleans, Melina Bell, Chris Macleod, Helen McCabe, Piers Norris Turner, and three reviewers.

- Matsuda, M. J., C. W. Lawrence III, R. Delgado, and K. Williams Crenshaw.** 1993. *Words That Wound: Critical Race Theory, Assaultive Speech, and the First Amendment* (Boulder: Westview).
- Mill, John Stuart.** 1969. *Utilitarianism*, in *Collected Works of John Stuart Mill*, vol. X, ed. by John M. Robson (Toronto: University of Toronto Press), pp. 203–59.
- Mill, John Stuart.** 1977. *On Liberty*, in *Collected Works of John Stuart Mill*, vol. XVIII, ed. by John M. Robson (Toronto: University of Toronto Press), pp. 213–310.
- Mill, John Stuart.** 1982. Notes on the Newspapers, in *Collected Works of John Stuart Mill*, vol. VI, ed. John M. Robson (Toronto: University of Toronto Press), pp. 149–280.
- Miller, Dale E.** 2010. *J. S. Mill: Moral, Social and Political Thought* (Cambridge: Polity).
- Nordenstreng, Kaarle.** 2016. Liberate Freedom from Its Ideological Baggage!, in *Freedom of Expression and Media in Transition: Studies and Reflections in the Digital Age*, ed. by Ulla Carlsson (Gothenburg: Nordicom), pp. 61–65.
- O'Rourke, K. C.** 2001. *John Stuart Mill and Freedom of Expression* (London: Routledge).
- Rawls, John.** 1971. *A Theory of Justice* (Cambridge, MA: Belknap, 1971).
- Rees, John.** 1960. A Re-reading of Mill on Liberty, *Political Studies*, 8: 113–29.
- Riley, Jonathan.** 1998. *Routledge Philosophy Guidebook to Mill on Liberty* (London: Routledge).
- Riley, Jonathan.** 2005. J. S. Mill's Doctrine of Freedom of Expression, *Utilitas*, 17: 147–79.
- Riley, Jonathan.** 2015a. *Routledge Guidebook to Mill's On Liberty* (New York: Routledge).
- Riley, Jonathan.** 2015b. Is Mill an Illiberal Utilitarian?, *Ethics*, 125: 781–96.
- Saunders, Ben.** 2016. Reformulating Mill's Harm Principle, *Mind*, 125: 1005–32.
- Skorupski, John.** 1989. *John Stuart Mill* (London: Routledge).
- Turner, Piers Norris.** 2013. Authority, Progress, and the “Assumption of Infallibility” in *On Liberty*, *Journal of the History of Philosophy*, 51: 93–117.
- Turner, Piers Norris.** 2014. “Harm” and Mill's Harm Principle, *Ethics*, 124: 299–326.