

Note: the final and definitive version of this article is published in *Legal Theory* 19 (03): 307–330, © Cambridge University Press 2013; available online via: [journals.cambridge.org/action/displayFulltext?type=1&fid=9152406&jid=LEG&volumeld=19&issueld=03&aid=9152403&fromPage=cupadmin&pdftype=6316268&repository=authInst](https://journals.cambridge.org/action/displayFulltext?type=1&fid=9152406&jid=LEG&volumeld=19&issueld=03&aid=9152403&fromPage=cupadmin&pdftype=6316268&repository=authInst)

## Intellectual agency and responsibility for belief in free speech theory

Robert Mark Simpson

Abstract: The idea that human beings are intellectually self-governing plays two roles in free speech theory. Firstly, this idea is frequently called upon as part of the justification for free speech. Secondly, it plays a role in guiding the translation of free speech principles into legal policy, by underwriting the ascriptive framework through which responsibility for certain kinds of speech harms can be ascribed. After mapping out these relations, I then ask what becomes of them once we acknowledge certain very general and profound limitations in people's capacity for intellectual self-governance. I argue that acknowledging these limitations drastically undermines the putative justifications for free speech of the type that I identify in the first part of the paper. I then show how we can reformulate an 'ascriptive' account of intellectual and doxastic responsibility, on which we may still be *held* responsible for what we think or believe, even though we lack agent-causal capacities in our thinking and believing. Then, in light of this ascriptive account of intellectual responsibility, I show how the key idea – that people 'have minds of their own' – can (and should) still play a significant role in guiding the legislative outworkings of free speech.

## 1. Introduction

My aim here is to map out the relationship between, on one hand, the idea that people have a capacity for intellectual self-governance – in colloquial terms, the idea that people ‘have minds of their own’ – and on the other hand, the complex schema of values and principles that standardly goes by the name ‘free speech’. This relationship is both interesting and important. Free speech occupies an exalted place in liberal democratic thinking, and the idea that people have minds of their own has shaped the way free speech is theorised and put into practice. Having said that, in the first part of this paper (§§2–5) I will suggest that philosophical work on free speech typically tells only a partial story about the significance of ideas about people’s intellectual self-governance, owing to a preoccupation with the justificatory foundations of free speech, and a lesser concern for the details of its policy outworkings. In the second part of the paper (§§6–8) I consider how we should think about the relationship between free speech and intellectual agency once we acknowledge the decidedly meagre extent of people’s intellectual agency. Some very influential defences of free speech seriously misrepresent the degree to which people govern their own thinking and believing. If we want to maintain that people are generally responsible for their beliefs, this claim cannot be grounded in a view of believing as a volitional undertaking, or as something akin to one. Here I present an alternative account of responsibility for belief, according to which we are responsible for our beliefs, in the sense of being answerable or accountable for them, because we ascribe – or under favourable reflective conditions we *would* ascribe – such responsibility to ourselves. I then explain how the idea that we have minds of our own, when it is understood from this kind of ascriptive perspective, can still play an important role in shaping free speech theory after all, but only via the relatively neglected connection that I identified in the first part of the paper, that is, in guiding the policy outworkings of free speech, rather than in being part of its justificatory foundations.

## 2. Free speech theory

Free speech is a mid-level political value. By this I mean that although it is the kind of value that one can invoke in order to defend specific policies, practices, institutions, etc., and to explain their aims and purposes, it is not the kind of value whose invocation brings us to justificatory bedrock. When we accept that some policy would serve free speech, it always remains a further question whether and why free speech should be privileged, such that its demands take precedence over the other ideals and desiderata

that are at stake in connection with a given policy.<sup>1</sup> I use the term ‘free speech theory’ to refer to the whole multifaceted normative schema that we attend to in thinking about such issues: (i) the axiological foundations upon which free speech as a mid-level political value rests, (ii) the policies, practices, and institutions that are putatively justifiable by appeal to this mid-level value, and (iii) the explanatory or justificatory relationships between foundations, the mid-level value, and policies.

Judges, lawyers, and legislatures are naturally preoccupied with the part of free speech theory that concerns the relationship between mid-level value and policy. In a given jurisdiction, J, there will be institutional parameters (e.g. provisions in J’s legal constitution, or articles in international treaties to which J is a signatory) that call for adherence to some more or less elaborated free speech ideal. The officials’ responsibility is to preside over J’s legal system so that it attains to this ideal, both in its *de jure* structure and its *de facto* practices.<sup>2</sup> Philosophers and political theorists, who are not subject to the onerous institutional duties that judges, lawyers, and legislators face, have instead tended to focus on the part of free speech theory that concerns the relationship between mid-level value and foundational axiology. On one hand we may see this as just a sensible division of labour: legal officials have expertise in one area, philosophers in another. But it also reflects something about the contemporary politics of free speech. By the late 20<sup>th</sup> century, liberal polities – the United States more than any other – had increasingly come to treat free speech as a preeminent political value. Lawmakers gravitated towards an approach on which, in the balancing of free speech with matters of civic harmony or public order, free speech presumptively took priority, so that only imminent and very serious threats to people’s welfare could justify speech restrictions. Political theorists and philosophers have, understandably, been absorbed by the question of

---

<sup>1</sup> In the literature this construal of the problem is often revealed through the author’s focussing on a particular case in which an allegiance to free speech is cited to justify the toleration of very harmful speech, in order to then ask what could be so important about free speech that it would warrant the toleration of such harm. A notable work that exemplifies this approach is Lee C. Bollinger’s *The Tolerant Society: Freedom of Speech and Extremist Speech in America* (Oxford: Clarendon Press, 1986); this work on free speech and tolerance was expressly conceived as a response to the Skokie Affair, in which the Illinois Supreme Court ruled in favour of the National Socialist Party of America’s right to parade through the town of Skokie while dressed in Nazi regalia.

<sup>2</sup> For recent work on free speech ideals, their position in international law, and their comparative place in various state jurisdictions, see for instance Ivan Hare, “Extreme speech under international and regional human rights standards”, in I. Hare and J. Weinstein (eds.), *Extreme Speech and Democracy* (Oxford: Oxford University Press, 2009); L. W. Sumner, “Incitement and the regulation of hate speech in Canada: a philosophical analysis”, in I. Hare and J. Weinstein (eds.), *Extreme Speech and Democracy* (Oxford: Oxford University Press, 2009); James Weinstein, “An overview of American free speech doctrine and its application to extreme speech”, in I. Hare and J. Weinstein (eds.), *Extreme Speech and Democracy* (Oxford: Oxford University Press, 2009).

why this should be: whether free speech *should* trump other things we rightly value.<sup>3</sup> There has been a focus on the justificatory questions, then, not only due to a division of labour, but because the orientation of modern liberal legal systems has brought the justificatory questions into stark relief.

One family of responses to the justificatory questions about free speech are what we might call ‘social good arguments’. Such arguments appeal to social goods that free speech promotes, such as a healthy political system, or excellence in education, or a vibrant and creative cultural milieu, or innovation in science and technology. Another family of responses to the justificatory questions are what we might call ‘individualistic arguments’. These arguments focus on the welfare, or interests, or dignity, or (putative) rights of individuals, and they purport to show how these things are promoted/honoured in permissive policies around speech, and diminished/dishonoured in censorious policies. Individualistic arguments emerge from a joint perspective, but they come with different points of emphasis. One argument focuses on the supposed inherent value of expression; it says, roughly, that as individuals with our own perspective on things, it is a good thing that we are allowed to share our views with others – in public speech, artistic expression, the written word, and so on. Another individualistic free speech argument, notably advanced by Frederick Schauer, proceeds from a libertarian concern regarding the potentially dangerous character of governments. According to this argument, governments have an invidious tendency to try to cement their own power, by repressing opposition and stifling dissent. Governments left to address free speech issues on a case-by-case basis will all too often get things wrong; they will downplay the benefits – and exaggerate the dangers – associated with confrontation, dissent, and social unrest. And hence, so the argument goes, societies need robust safeguards for free speech.<sup>4</sup>

---

<sup>3</sup> In the relevant jurisprudential literature, this question receives its most perspicuous formulation and analysis in Frederick Schauer, “Must speech be special?”, *Northwestern University Law Review* 78 (1983): 1284-1306.

<sup>4</sup> See in particular Frederick Schauer’s ‘argument from government incompetence’ in *Free Speech: A Philosophical Inquiry* (Cambridge: Cambridge University Press, 1982), pp. 86ff. An especially strident example of this type of argument comes in Vincent Blasi’s paper on what he calls ‘the pathological perspective’ on free speech; “the overriding objective at all times” Blasi says, “should be to equip the first amendment to do maximum service in those historical periods when intolerance of unorthodox ideas is most prevalent and when governments are most able and most likely to stifle dissent systematically”; Vincent Blasi, “The pathological perspective and the First Amendment”, *Columbia Law Review* 85 (1985): 449-514, pp. 449-50.

### 3. Intellectual agency as a justificatory basis for free speech

Here I want to home in on a third type of individualistic argument, one that has been prominent in the work of liberal political theorists and philosophers, and in the jurisprudence they influence. In this type of argument, an individualistic approach to free speech theory is adjoined to a particular thesis about human beings' capacities, roughly, the thesis that people are in some significant sense intellectually self-governing. In colloquial terms, the view is that people 'have minds of their own'. For free speech advocates who advance this kind of argument, the guiding vision of a society in which people are assured of the freedom to share their opinions with others is justified by appeal to a corollary notion: that we all have the capacity to make up our own minds about things and to think for ourselves. According to proponents of this view, then, it is wrong to restrict speech that aims to foment unrest, or to stir enmities against individuals, groups, or governments. Why? Because we have the capacity to think for ourselves when exposed to speech of this sort, and we must be trusted to engage with such speech accordingly.

Some explanatory work is needed to fill in the details of how an appeal to the thesis about our intellectual agency lends support to free speech policies. Probably the most well-known way of spelling out this justificatory relation comes from J. S. Mill's discussion in *On Liberty* (1859). In short, Mill regards the personal development of intellectually autonomous human individuals as something of the utmost value, and he sees free speech as playing an essential role in fostering the conditions under which such development is able to take place. Mill is an archetypal individualist. He regards the proper domain of human liberty as "that portion of a person's life which affects only himself or, if it also affects others, only with their free, voluntary, and undeceived consent and participation".<sup>5</sup> He insists, therefore, that a person should be unconstrained in the things that only *immediately* affect herself: what she thinks, feels, and believes. But notice that speech is unlike these other inwardly-directed mental goings-on. Speech *can* immediately affect others without their consent. So why does Mill think we must also be free to express our thoughts, feelings, and beliefs? In short, because he understands free speech as being 'practically inseparable' from liberty of thought and conscience *per se*, and we need liberty of thought and conscience – and by extension, free speech – because without them people will not be able to attain to the kind of creativity, inventiveness, and originality in their lives, from which, as Mill puts it, "the initiation of all wise or noble things comes".<sup>6</sup> Utility is ultimately the name of the game for Mill, but it is utility "in the largest

---

<sup>5</sup> John Stuart Mill, *On Liberty*, ed. Gertrude Himmelfarb (Harmondsworth: Penguin, 1985 (originally published 1859)), p. 71.

<sup>6</sup> *Ibid*, p. 131

sense, grounded on the permanent interests of man as a progressive being”.<sup>7</sup> To promote this sort of utility we need to foster an environment which allows individuals to engage in fearless inquiry, innovation, and even eccentricity, and this in turn means we must protect liberty in thought, conscience and speech. The alternative is to live in a world where people are hemmed in by the tyranny of the majority, and where everyone’s lives are attenuated by a culture of dogmatism and conformity.

Another – less widely-known, but related and complementary – view of the relationship between intellectual agency and free speech can be found in the work of contemporary political theorists like Thomas Scanlon and Ronald Dworkin. According to these authors governments must refrain from restricting dangerous or unpopular speech, not because of a consequentialist imperative to cultivate conditions in which our intellectual capacities can flourish, but because a government relinquishes its claim to *legitimate* authority over a society of intellectually independent agents, if and when its laws do the indignity to those agents of not respecting their capacity to think for themselves. Proponents of this view accept that unfettered speech may foment harm and turmoil. They accept that some who are exposed to dissent, hostile invective, or provocation, will be influenced towards harmful behaviour. What proponents of this view say, however, is that even while recognising these risks, no-one who regards herself as a fully autonomous and rational agent can reasonably consent to living under the rule of a government which prevents her from engaging with radical, or outré, or controversial opinions. And if this is right – if such agents wouldn’t consent, under fair conditions, to honouring the demands of a regime which carries out this kind of censorship – then the censorious regime is, *ipso facto*, an illegitimate regime.<sup>8</sup> Dworkin spells out this line of reasoning in the following way; “morally responsible people”, he says

insist on making up their own minds about what is good and bad in life or politics, or what is true and false in matters of justice or faith. Government insults its citizens, and denies their moral

---

<sup>7</sup> *Ibid*, p. 70

<sup>8</sup> The canonical version of this argument is in Thomas Scanlon, “A theory of freedom of expression”, *Philosophy & Public Affairs* 1 (1972): 204-26. In subsequent work responding to critics – particularly Robert Amdur, “Scanlon on freedom of expression”, *Philosophy & Public Affairs* 9 (1980): 287-300 – Scanlon backed away from the hard-line formulation of this argument, conceding that it is merely our interest in autonomy as a good to be promoted, rather than autonomy as some sort of inalienable dignitarian right, which grounds the obligations of governments to refrain from restricting things like radical dissident speech; see Thomas Scanlon, “Freedom of expression and categories of expression”, *University of Pittsburgh Law Review* 40 (1978): 519-50. To be clear, when I speak of ‘Scanlon’s account of free speech’ in what follows here, I’m referring to the position espoused in the more well-known 1972 paper, which is not to be seen as Scanlon’s current or final or official view.

responsibility, when it decrees that they cannot be trusted to hear opinions that might persuade them to dangerous or offensive convictions.<sup>9</sup>

There are differences in the normative theoretical perspectives of the two lines of argument sketched here. Mill is operating within a consequentialist framework. Dworkin and Scanlon employ a deontological contractualist framework, with an emphasis on the conditions of legitimate governmental authority. The common ground that I want to stress, though, is that both views advert to people having minds of their own as part of the underlying justification for a regime of free speech. When we are trying to explain why free speech should occupy a privileged position in our constellation of political values, these views identify human beings' intellectual agency, along with some adjoining ethical principles which explicate the normative implications of that agency, among the principal justifying considerations.

#### **4. Responsibility-attribution in free speech policy**

It is important to remind ourselves that theorists in the vein of Mill, Scanlon and Dworkin do not seek to defend an *absolutist* free speech regime, in which it is always illegitimate to restrict free speech for the sake of other values. Their claim is that there are weighty reasons – including reasons grounded in ideas about people's intellectual agency – for having a legal system in which free speech enjoys robust institutional safeguards. But advocates of these views readily acknowledge that certain restrictions on speech are sometimes required, e.g. where they are needed to prevent the infliction of egregious personal harms, or to avert a major disaster. One of the perennial questions for defenders of free speech, then, is how these sorts of trade-offs should be negotiated. The question is difficult because, whatever normative theoretical framework we are operating within, decisions about how to balance free speech against other values and desiderata appear to hinge upon our answer to un-settleable questions about our axiological priorities; questions like: just how important *is* free speech, relative to other values and desiderata? Or: how much inequality, or danger, or humiliation *should* we tolerate for the sake of free speech?

I say that these questions are un-settleable, but of course they do receive a settlement, whether explicit or implicit, in the policies that our governments implement in order to regulate harmful and dangerous speech. What I want to claim here is that philosophers and political theorists have tended to

---

<sup>9</sup> Ronald Dworkin, "Why must speech be free?" in *Freedom's Law: The Moral Reading of the American Constitution* (Cambridge Massachusetts: Harvard University Press, 1996), p. 200. For a more recent re-statement of Dworkin's argument in this vein, see Dworkin, "Foreword" in I. Hare and J. Weinstein (eds.), *Extreme Speech and Democracy* (Oxford: Oxford University Press, 2009).

overlook the way in which ideas about people's intellectual agency fulfil an important purpose in this part of free speech theory: the part that comes after we've granted that free speech is an important value, and when we are trying to decide just how valuable it is, and by extension, how it should be balanced with other values and desiderata.<sup>10</sup> My claim is that the idea that 'people have minds of their own' matters greatly in this part of free speech theory, because it guides judgements about when speakers may be made accountable for adverse outcomes to which their speech causally contributes.

Let's start with a very simple toy example. Suppose we are in a setting in which, in the local vernacular, the phrase "X ought to be shot" is conventionally taken to mean that X is a foolish and annoying person. Now imagine a case in which A says to B that C ought to be shot, and in which B, to A's horror, interprets these words literally and shoots C. In this case it is true that A causes C to suffer an injury, in the bare counterfactual sense that if A had not spoken to B, C would not have been shot. From one perspective, then, the question of whether to restrict utterances like A's utterance – in this and other structurally similar scenarios – may be construed as a question about our axiological priorities: should we prioritise freedom in the use of mildly coarse vernacular language, over the legitimate aim of protecting people against verbally-provoked violence? Now, most readers will (I expect) favour expressive liberty in this kind of scenario. But so far as this is the case, it would be a mistake to interpret this as the staking of an axiological claim to the effect that free speech is more valuable or important than the prevention of miscommunication-inspired violence. Instead, we should understand this sort of case as pivoting on a judgement about responsibility-ascription, for in so doing we can offer an answer to the policy question without pretending to settle any questions of axiological priority. In the scenario at hand, let's assume that B normally possesses, and is (reasonably and correctly) supposed by others to possess, a typical measure of intellectual agency – say, what we would informally describe as the capacity to think for herself. It follows, in that event, given plausible mediating precepts about the relationship between causation, control, and responsibility, that B's act of shooting the weapon, rather than A's utterance, is where proximate *causal* responsibility for C's injury lies, and where responsibility in the sense of *accountability* should be allocated in the first instance. We need not agonise about a conflict between free speech and

---

<sup>10</sup> The most notable exception that I know of, in stressing the importance of responsibility-ascriptions in resolving apparent axiological conflicts around free speech policy, is Evan Simpson, "Responsibilities for hateful speech", *Legal Theory* 12 (2006): 157-77. Rather more attention is paid to this connection in the jurisprudential literature on incitement law, and unsurprisingly so, since the central issue in that area of law is the conditions under which responsibility redounds to the inciting or the incited party; see for instance David Strauss, "Persuasion, Autonomy, and Freedom of Expression", *Columbia Law Review* 91 (1991): 334-71; Larry Alexander, "Incitement and freedom of speech", in L. Alexander (ed.) *Freedom of Speech* (Aldershot: Ashgate, 2000).



the prevention of violence, on this view, because neither the speaker nor her speech are responsible for the violence that B inflicts on C. Any legislative undertaking which sought to restrict speech in these sorts of cases – whatever the legislators’ intentions might be, and whatever outcomes might eventuate – would not in fact serve to prioritise the protection of people against violence over free speech. Rather, it would enshrine an erroneous judgement about where responsibility lies for harms brought about by certain kinds of actions whose performance is instigated by other people’s verbal conduct.

This point about responsibility-ascription naturally consorts with what is sometimes called ‘the persuasion principle’ in free speech theory, according to which “the government may not justify a measure restricting speech by invoking harmful consequences that are caused by the persuasiveness of the speech”.<sup>11</sup> The persuasion principle may be cited to justify legal protection for those who speak in favour of violent revolution, say, or theories of racial superiority. Of course it is true that when person B comes to believe that he should commit acts of violence or racist discrimination, harm may eventuate. And of course it is entirely possible that people who engage in these forms of speech may influence B towards such beliefs. But if B is fully responsible for his beliefs, then even if the speech of others does contribute to his harmful action, by exerting an influence on his beliefs, the speakers should not incur liability for the harm that B’s action inflicts. Or so the persuasion principle tells us. The common thread between the persuasion case and the miscommunication case can be spelled out as follows. In each case person A’s speech sets off a causal sequence which results in harm being done to another. But in neither case is A responsible for that harm, because the speech effects its harm by eliciting a response from another party, who is, or who can be justifiably treated as being, responsible for his own action, despite the fact that the action is influenced by, and would not have occurred but for, A’s initial speech.

It is through this kind of connection that our key idea – that people have minds of their own – carries out its further role of guiding the translation of free speech values into free speech policy. Claims about our intellectual agency can be invoked as part of the general justification of why we should privilege speech in the first place. This is what we observe in the individualistic free speech arguments from Mill, Scanlon, and Dworkin. My point is that such claims are also at work in determinations about how the law allocates *causal* responsibility – and thus underwrites its *ascriptions* of responsibility, in the sense of blame, liability, or culpability – for certain kinds of harmful outcomes associated with speech.

---

<sup>11</sup> Strauss, *supra* note 10, p. 334.

## 5. Complications in the simple picture of responsibility-ascription

What I've presented so far is a highly simplified picture of how such responsibility-ascriptions operate. At least two sources of complexity must be incorporated into this picture. First, we must acknowledge that responsibility (in the sense of blame, liability, or culpability) comes in degrees. Person B might be more culpable for being *persuaded* to act violently than she would be for being *provoked* into violent action. But it is consistent with this – and independently plausible, besides – to say that B is culpable to some degree in both cases. Second, and relatedly, we must acknowledge that responsibility can be incurred by both – or as the case may be, by *all* – participants in a sequence of events that unfolds along the kind of schematic trajectory that we've been considering here, in which (i) *speech* exerts (ii) an *influence*, which leads to (iii) an *action* that results in (iv) *harm*. That is to say, even where events can be mapped out in terms as clear and linear as this, there will still be cases in which both the speaker, A, and the person who reacts to A's speech, B, can incur some measure of culpability for harm that results from B's action. Incitements and provocations to violence are two kinds of cases in which this relation plausibly holds.<sup>12</sup>

So how does responsibility ascription in this arena work, once these complications are taken into account? And most importantly for our purposes, how is the idea that people have minds of their own involved in this? Communicative acts are part of the causal backstory for all sorts of harmful outcomes. Sometimes the causal connections are ostensibly linear, as in the simple sequences sketched above. In other cases, though, causality is more diffuse. For instance, there may be many people whose speech makes a small contribution to a social climate which gives rise to actions or patterns of behaviour that inflict harm on people. When the harms are finally inflicted, the individuals who perform the proximate harm-causing actions are presumptively answerable for those harms. But the extent of their answerability depends, among other things, on the form of the communicative influence in whose wake they were acting. Someone who is incited to violence by another person's speech is, other things being equal, not answerable to the same degree as someone who is persuaded to act violently, or who acts violently on a mere suggestion. The mode of communicative influence also matters in determining to what extent, if at all, responsibility redounds to speakers who causally contribute to the preconditions which encourage or give rise to harmful actions. The speaker who goads, or provokes, or verbally bullies someone into acting

---

<sup>12</sup> What is being rejected here is what we may call (borrowing some terminology from Robert Nozick) a 'bucket' theory of responsibility. On this view there is a fixed quantum of liability,  $\alpha$ , associated with a given harm,  $x$ , to be allotted between A and B, such that if A is solely liable for  $x$ , A incurs exactly  $\alpha$  units of liability, or, if A and B are equally liable for  $x$ , they both incur  $0.5\alpha$  units. See Robert Nozick, *Anarchy, State, and Utopia* (New York: Basic Books, 1974).

badly, will be more culpable for harms that eventuate from his speech than the speaker who makes a remark which unwittingly sets off an unforeseen chain of events. Our ascriptions of liability in all these sorts of cases will also depend upon further considerations besides the nature of the communicative interaction between speaker and audience. We must consider the duties that people have by virtue of the social roles they occupy. We also must consider the intentions and expectations that people manifest in their communicative interactions, and in the actions that follow them. And crucially, we must figure out how to negotiate this, and lay the liabilities at the right pairs of feet, in a manner that balances free speech ideals, on one hand, with the aims of harm-prevention and -deterrence on the other.

There are two ways in which the idea that people have minds of their own shapes and organises the responsibility-ascription framework that operates in this policy arena. Firstly, in cases where speech leads to harm via an ostensibly linear sequence of events, this idea can be invoked – as one factor among others – in order to guide the rendering of specific verdicts about which participant or participants are legally answerable for specific harms, and to what extent. Secondly, this idea can be called upon in order to explain or defend the whole diagnostic undertaking of trying, in the first place, to distinguish between the cases that involve (or are akin to) persuasion, and those that involve (or are akin to) provocation or incitement. If we *did* think of people as entirely suggestible or malleable in response to the speech of others, there would be little sense to be made of the law's seeking to make these discriminations. It is primarily *because* we see people as having minds of their own, that we regard quite different allocations of culpability and blame as being called for in, say, (i) the case where the speaker, A, consciously attempts, via acts of incitement or provocation, to bypass the listener B's capacity to act under his own reflective volition, and (ii) the case in which A's speech contributes (via subtle influence) to an act performed by B which was, by A's lights, unforeseen, or unintended, or just beyond her (A's) concern.

We can see these considerations at work in complex real-world policy problems which involve harmful outcomes associated with expressive conduct. Take the case of pornography. Feminist scholars like Catharine MacKinnon, Andrea Dworkin, Rae Langton, and Jennifer Hornsby argue that pornography subordinates and silences women, by legitimating demeaning social attitudes towards women.<sup>13</sup> Given this claim, we can understand legislative resolutions vis-à-vis the restriction of pornography as hinging on the

---

<sup>13</sup> Some of the key works in this connection include Andrea Dworkin and Catharine A. MacKinnon, *Pornography and Civil Rights: A New Day for Women's Equality* (Minneapolis: Organizing Against Pornography, 1998); Rae Langton, "Speech acts and unspeakable acts", *Philosophy & Public Affairs* 22 (1993): 293-330; Catharine A. MacKinnon, *Only Words* (London: HarperCollins, 1994); Jennifer Hornsby, "Free speech and illocution", *Legal Theory* 4 (1998): 21-37.

question of whether we should prioritise expressive liberty (in the domain of sexually explicit words and images) over sexual equality. Decisions about anti-pornography law, thus construed, weigh free speech against aspirations to sexual equality. But there is an alternative construal. We can think about this issue in a way that stresses responsibility-ascription, rather than seeing it as a showdown between two overarching ideals, and when we do the conflict looks rather different. Let's assume that pornographic materials on the whole make *some* sort of contribution to the wide-scale, multifaceted systems of sexual inequality that are at work in contemporary societies. The contribution in question (whatever its *extent* may be) will be brought into effect, principally, via the exertion of a malign influence upon the beliefs and attitudes of those who consume, or who are in some other way met with, pornographic representations of women. Given this way of framing things, the question of whether to legally restrict pornography hinges above all on whether, and if so to what degree, those who produce and distribute pornography can be adjudged responsible for the later (harmful) outcomes that ensue in the wake of the influence pornography has upon its audience. If we think of pornography's influence as being akin to persuasion, then we will judge that the persuaded parties are relatively more culpable, and the persuading parties relatively less culpable (if at all), for the harms in question. If we think of pornography's influence as being akin to incitement or conditioning, then we will judge that the incited parties are less culpable, and the inciting parties more culpable, for such harms.<sup>14</sup> Findings about the extent to which people exercise intellectual agency in this arena (for instance: if and how consumers of pornography subconsciously or involuntarily change their judgements about women after exposure to pornography) are crucial in figuring out where, along this spectrum, the communicative interactions in question should be situated. And our initial notion – that people *can* exercise a greater or lesser degree of intellectual agency – is crucial in making sense of why we would recognise such a spectrum, and its normative ramifications, in the first place.

---

<sup>14</sup> The anti-pornography argument may be formulated (as in, for example, Rae Langton, "Subordination, silence, and pornography's authority", in R. C. Post (ed.) *Censorship and Silencing: Practices of Cultural Regulation* (Los Angeles: Getty Research Institute for the History of Art and the Humanities, 1998)) such that the responsibility that pornographers are attributed in relation to the perpetuation of sexual hierarchy is based not merely upon the allegation that pornography influences people towards sexist beliefs and attitudes, but on the claim that pornographers have an *authoritative* voice in the social conversation about sexual relations. As Leslie Green argues, though, this claim about pornography's authority is not easily reconcilable with the way that pornography is, despite its widespread use in people's private lives, socially coded as tawdry, low-class speech, nor with the fact that other competing sources of putative norms concerning sexual relations (e.g. in law, in academia, in the education system, and in religion) do carry some degree of moral and intellectual authority – more than is carried by pornography, at any rate. See Leslie Green, "Pornographizing, subordinating, and silencing" in R. C. Post (ed.) *Censorship and Silencing: Practices of Cultural Regulation* (Los Angeles: Getty Research Institute for the History of Art and the Humanities, 1998).

## 6. Involuntarism and doxastic responsibility

To summarise, the idea that people have minds of their own plays two roles in free speech theory. Firstly, it is called upon to lend justificatory support to special protections for speech. And then, secondly, it helps to shape and organise the ascription of responsibility for speech-related harm.

So much for how we theorise the relations in question; the problem with all this, so I now want to suggest, is that our central idea – that we have minds of our own – is implausible, or to put it more cautiously, the understandings of that idea which would straightforwardly make sense of the way the idea is used in free speech theory are implausible. There are several routes through which we can start to pull apart the notion that we as human beings have minds of our own. We might, for instance, advert to the mass of empirical evidence about people's tendency towards doxastic conformity, and groupthink, and herd psychology.<sup>15</sup> Or we might consider the many respects in which our beliefs, both in their specific contents and in their broader architecture, seem to be governed by accidents of geography.<sup>16</sup> Or we might look to the findings of situationist psychology – in particular, those in which the framing considerations or priming factors in the subject's environment are shown to exert a large and immediate influence upon the subject's beliefs, thoughts, and judgements.<sup>17</sup> In each case we will find a significant challenge to the idea of our intellectual individuality. We will find reasons to think of a person's doxastic constitution as a product of her environment, rather than as an expression of some governing mental self. But there is, in addition to these kinds of challenges, a more elemental misapprehension in the widely-received view of human beings as intellectually agential beings. The problem with this view, to put it bluntly, is that we are simply unable to decide upon or govern what we believe. To have a mind of one's own is, so one may naturally suppose, to have authorship or control over one's important mental goings-on, in something like the way one has authorship over properly volitional acts. But believing that *p* is not a volitional act, like assenting to *p*, or asserting that *p*. Believing that *p* is just something that happens to the subject in light of whatever *p*-relevant considerations have a purchase on her routine (largely automatic and unconscious)

---

<sup>15</sup> For reviews of the groupthink literature see for instance James K. Esser, "Alive and well after 25 years: A review of groupthink research", *Organizational Behavior and Human Decision Processes* 73 (1998): 116-41; Robert S. Baron, "So right it's wrong: groupthink and the ubiquitous nature of polarized group decision making", *Advances in Experimental Social Psychology* 37 (2005): 219-53.

<sup>16</sup> A wealth of evidence pertaining to which is helpfully summarised in Richard E. Nisbett, *The Geography of Thought: How Asians and Westerners Think Differently – And Why* (New York: Free Press, 2003). For a more general discussion around this issue see also Jesse Prinz, "Culture and Cognitive Science" in E. N. Zalta (Ed.) *The Stanford Encyclopedia of Philosophy* (Winter 2011 Edition), accessible via <http://plato.stanford.edu/archives/win2011/entries/culture-cogsci/>.

<sup>17</sup> See for instance John M. Doris, *Lack of Character: Personality and Moral Behavior* (New York: Cambridge University Press, 2002).

processes of cognition. This claim is not supposed to rest on any bold metaphysical theses about our freedom in general or lack thereof. The suggestion is not that we are metaphysically unfree, and therefore unfree as believers. Rather, the suggestion is that even if most people can control or govern some significant range of their actions, in a way that is sufficient to constitute their being metaphysically free *per se*, this agential capacity does not extend into the doxastic realm. A subject's belief that *p*, or her credence regarding *p*, is something she finds herself having, rather than something she does.

One way to understand why this is the case, which I take from Bernard Williams's well-known work on this subject, is to conceive of beliefs as inward representations of how one takes the world to be. The basic explanation for the phenomenon, then, is that we cannot decide to believe such-and-such, because we cannot, by force of volition, take the world to be a different way from how we *actually* – that is, how we artlessly and ingenuously – take it to be.<sup>18</sup> Now, when we say that people have a capacity to think for themselves, or something along these lines, the easiest way to make sense of that claim is to consider the contrast that it implies. Some people are doxastically malleable. Consider a very young child, or a victim of brainwashing, or an extremely pathologically gullible person: these people have a special kind of lack in authorial control over their mental lives. As doxastic subjects, they will be completely at the mercy of the circumstances in which they find themselves. The picture of our intellectual agency that is invoked in arguments like those we find in Mill, Scanlon, and Dworkin, naturally contrasts with such cases. To say that we have minds our own is to say that, unlike these folks, *our* beliefs are not helplessly determined by circumstances outside our control. Having a mind of one's own, whatever else it involves, at least means having a capacity to govern oneself in what one thinks and believes. The worry, then, is that it is difficult to see how this capacity can be understood as residing in any of us, once we recognise the involuntaristic character of belief. Because all of us, along with the naïf and the brainwashed person, simply find ourselves having certain beliefs, and taking the world to be a certain way. It is true that the naïf and the brainwashed person are beholden to the deliberate manipulations of other people, in a way that most of us are not, most of the time. However, our not being beholden in that one respect is not enough to constitute our having an agential doxastic capacity. We may not be beholden to a manipulating other, but we are still beholden to the way we take the world to be, and the beliefs that represent this still come to us unbidden, and the whole process is still something that ultimately *befalls* each of us.

---

<sup>18</sup> Bernard Williams, "Deciding to believe" in *Problems of the Self: Philosophical Papers 1956-1972* (Cambridge: Cambridge University Press, 1973).

Let's call the thesis in question 'involuntarism'. The central pillar of the thesis is that people are in normal cases unable to decide upon their beliefs. To this we might also want to add, secondarily (and more tentatively), that people are similarly unable to exert volitional- or decision-making-style control over a large portion of their other mental states.<sup>19</sup> A more completely spelled-out version of the thesis will require some finessing and qualification. For one thing, it will have to take account of the fact that people can (and routinely do) act in ways that indirectly guide their doxastic attitudes. A long time ago Pascal, in his discussion of the Wager, suggested that if a person makes himself earnestly participate in a religious way of life, he can cause himself to gravitate towards the right (theistic) beliefs. If something like this is right, and if it is also true that we have agential capacities in decisions about religious participation etc., then it follows that we have some indirect agential control over some of our beliefs.<sup>20</sup>

Even thus qualified, however, involuntarism poses a serious problem for free speech theory's primary way of conceptualising the relationship between intellectual agency and free speech as a mid-level value. The justificatory view of that relationship, recall, says (i) that governments must allow free speech in order to foster the free-thinking intellectual individualism that human flourishing depends on (as in Mill), and/or (ii) that governments must not, on pain of illegitimacy, legislate in a manner that insults people's dignity as beings with minds of their own (as in Scanlon and Dworkin). When faced with the challenge of involuntarism, the conception of human agency to which those justifications appeal seems untenable. We are not free-thinking beings with minds of our own, or if we are, it is only in a relatively meagre, indirect sense. We can draw useful distinctions regarding the varieties and sources of doxastic unfreedom. The brainwashed person faces an additional and pernicious form of doxastic unfreedom that the rest of us do not. But none of us occupies a higher-order intellectual control centre, from which we choose how we will take the world to be. And if someone somehow acquired such an ability – an ability

---

<sup>19</sup> The secondary (more tentative) suggestion, then, is that a considerable portion of our non-doxastic attitudes – for instance: the inchoate musings that flit through our minds, some of our aesthetic judgements, and some of our preferences and desires – are belief-like in respect of their involuntaristic nature. We may allow that people have *some* reflective control over the content of such thoughts (for example, one can tell oneself to 'think happy thoughts'), but still *many* of these thoughts simply befall us, like our beliefs. It is of course consistent with this suggestion to say that people have an agential capacity to decide how they will behave *given that* they believe p, or intend x, or find y appealing, or are upset by z. The idea is just that in the usual run of cases people do not decide to adopt, nor effectively deliberate about whether to maintain, these arrays of mental states.

<sup>20</sup> See Sharon Ryan's entry on doxastic voluntarism (pp. 322-325) in J. Dancy, E. Sosa, and M. Steup (Eds), *A Companion to Epistemology* (Chichester: Wiley-Blackwell, 2010). For an attempt to ground a general account of responsibility for belief in claims about our capacity for *indirect* doxastic self-control, see Mark Leon, "Responsible believers", *The Monist* 85 (2002): 421-35.

to cajole herself into seeing the world however she chose to see it – then so much the worse for her, and for the prospect of people’s intellectual agency if that is what it essentially amounts to.<sup>21</sup>

The appeal to claims about our intellectual autonomy as part of the putative justification for free speech has been criticised by others in the contemporary philosophical literature. Susan J. Brison, for example, highlights the way in which a conception of autonomy as “the actual ability to be rationally self-legislating” is invoked to defend orthodoxy free speech politics. And in response to this, she argues that such free speech policies are in fact much less well-equipped for promoting people’s autonomy (in the relevant sense) than many of the reformist (and pro-regulation) views of free speech that tend to draw the ire of free speech advocates.<sup>22</sup> In a similar vein, Susan Hurley discusses empirical evidence which reveals human beings’ considerable susceptibility to the imitation of violence depicted in the media, and shows how this evidence undermines claims about our intellectual autonomy and its normative ramifications in relation to free speech.<sup>23</sup> More generally, feminist critiques of mainstream free speech theory have often stressed how there is a picture of human agency underlying free speech justifications – a view of human agency as something we all have, equally, and inalienably – which is blind to the effects of social power and identity-based inequality, both in creating and in diminishing people’s agential capacities.<sup>24</sup> These authors are quite right to challenge free speech theory’s appeal to claims about our intellectual agency. What I’m suggesting here is that they do not go far enough. Defences of free speech which appeal to our intellectual agency rest on a false picture of our human nature; to reiterate, we are not free-thinking beings with minds of our own, and if we are, it is only in a relatively meagre and indirect sense.

---

<sup>21</sup> I should acknowledge that there are defences of doxastic voluntarism in the contemporary literature, e.g. Sharon Ryan, “Doxastic compatibilism and the ethics of belief”, *Philosophical Studies* 114 (2003): 47-79; Matthias Steup, “Doxastic freedom”, *Synthese* 161 (2008): 375-92. Note, however, that the general tenor of these pieces is compatibilistic. Their aim is to recover some sense in which we can be said to be doxastically free despite our lacking deliberative control over doxastic mental processes. Such efforts are broadly consistent with my contentions here, except that where my compatibilist undertaking is concerned to establish the justifiability of making doxastic subjects accountable for their beliefs *despite* their lacking agent-causal responsibility for those beliefs, others are concerned to articulate an account of these interconnected dimensions of doxastic responsibility in a manner such that the whole schema can be characterised as one on which the subject is in some sense doxastically ‘free’.

<sup>22</sup> See pp. 330ff. in Susan J. Brison, “The autonomy defense of free speech”, *Ethics* 108 (1998): 312-39.

<sup>23</sup> Susan Hurley, “Imitation, media violence, and freedom of speech”, *Philosophical Studies* 117 (2004): 165-218.

<sup>24</sup> Several recent examples of this critical approach can be found in Ishani Maitra and Mary Kate McGowan (Eds), *Speech & Harm: Controversies Over Free Speech* (Oxford: Oxford University Press, 2012).



## 7. Intellectual accountability without intellectual agency

Given all this, can we still make anything of the idea that we have minds of our own, in the sense of our being responsible for what we believe? And is there any way for this to still play a role in our theorising about free speech? In what remains I want to suggest, firstly, that even if involuntarism (or some duly-qualified version of the thesis) is correct, it can still be reasonable, proper, and justifiable to *hold* people responsible, in the sense of regarding people as accountable or liable, for believing as they do; and secondly, that by characterising the nature and grounds of intellectual responsibility in this way, we can re-establish some connections through which the idea of our intellectual agency will still have a role to play in free speech theory after all. Let's begin by drawing some distinctions between the related senses of 'responsibility' that can become commingled in an analysis of such issues. Following H. L. A. Hart,<sup>25</sup> we can distinguish three different senses of responsibility in law, under the following headings:

*Causal responsibility*: in relation to which to say that A is responsible with respect to some state of affairs,  $\alpha$ , is to say that from the law's perspective, A caused or was the principal causal contributor to  $\alpha$  (as in when we say that the vandal was responsible for the broken letterbox).

*Capacity-responsibility*: in relation to which to say that A is responsible vis-à-vis  $\alpha$  means that so far as A is implicated in  $\alpha$ 's occurrence, A had a capacity to (i) understand the normative demands to which she was subject vis-à-vis  $\alpha$ , (ii) to deliberate about her actions accordingly, and (iii) to control her  $\alpha$ -related actions as guided by her deliberations (as in when we say that the sleepwalker was *not* responsible for setting off the fire alarm).<sup>26</sup>

*Liability-responsibility*: in relation to which to say that A is responsible vis-à-vis  $\alpha$  means that A is liable to be punished or to make reparations in relation to some negative state of affairs in  $\alpha$  (as in when we say the energy corporation is responsible for rehabilitating the oil-ravaged coastline).

---

<sup>25</sup> H. L. A. Hart, "Postscript: responsibility and retribution" in *Punishment and Responsibility: Essays in the Philosophy of Law*, 2nd Edn (Oxford: Oxford University Press, 2008), pp. 214ff. Hart also identifies a fourth sense of responsibility which is less relevant to our discussion here – namely, responsibility as *role-based duty* – in relation to which to say that A is responsible with respect to an action  $\varphi$  is to say that A has a duty to  $\varphi$  because of a social role that she occupies, or a particular commitment that she has made to others (e.g. as in when we say that the surgeon has a responsibility to gain the patient's consent before operating).

<sup>26</sup> In such cases we are not denying that the sleepwalker was causally 'to blame' for the outcome. Rather, we are judging that the sleepwalker is not evaluatively blameworthy for the outcome, insofar as the nature of her causal connection to the outcome is not such that she stands to be evaluated on account of her causal role in the outcome; see Hart, "Postscript", pp. 227ff.

The standard way to conceive of the relationship between these three different senses of responsibility is to regard causal responsibility and capacity-responsibility as necessary conditions for the ascription of liability-responsibility. We should not punish bystanders, nor those whose incapacities make them immune to culpability in general, and this means that person A cannot be considered a candidate for bearing liability-responsibility in relation to an adverse state of affairs,  $\alpha$ , unless A at least bears causal responsibility for  $\alpha$  and possesses capacity-responsibility in relation to  $\alpha$ . Or so the familiar line of thought goes. Note, however, that even if the relevant conditions hold in a wide range of cases, there are relatively uncontroversial exceptions both in law and in everyday life. In everyday life parents punish children who are not yet capacity-wise responsible agents; we consider our neighbours accountable if their dog (or their fallen tree) damages our shared fence; I may expect my friend to compensate me for my broken vase, even though his breaking of the vase came from an involuntary flailing of his limbs upon seeing a spider. In such cases we ascribe blame and/or accountability to agents who lack either causal responsibility or capacity-responsibility (or both) with respect to the adverse outcome from which these negative liabilities stem. In law we punish various forms of criminal negligence, and we routinely allow penalties to be levied in private law (under the auspices of strict liability torts) to compensate for damages in relation to which the penalised party lacked proximate causal responsibility. Our commitment to these sorts of practices of *holding* people responsible, in both legal contexts and informal settings, gives expression to an important judgement: that under a certain range of circumstances, liability-responsibilities can properly redound to people independently of prior ascriptions of causal responsibility and capacity-responsibility.

Should we regard the domain of belief as another arena in which these patterns of judgement may be legitimately applied? Let's consider some examples. Suppose person A is an adult with a typical measure of intellectual independence – that is to say, A is not a victim of brainwashing or hypnosis, nor is he someone with the mental capacities of a child – and suppose A comes to believe that women are inferior to men, and that they deserve to be treated accordingly in the domain of sexual conduct. Does our acknowledgement that A's beliefs are involuntary in their underlying character, in the same way that everyone's beliefs are involuntary, deter us from our inclination to condemn A for his beliefs about sex? Or consider a case with the opposite evaluative valence. If person B grows up in a racist household but later rejects her belief in the inferiority of other races, don't we want to say that she merits praise for this, and isn't it the case that our inclination to praise B persists even once we remind ourselves that B's beliefs come to her unbidden, as for the rest of us? A few quick examples cannot bear any large argumentative

burden, but I think we can explain why the evaluative verdicts that I'm implicitly endorsing here may be acceptable after all, and need not be based on confusion or folk-psychological dogma.

With respect to involuntarily acquired beliefs and attitudes, normal, mature human agents *are* liable to praise, blame, reward, punishment, and so on, and we are liable in this way because we ascribe – or under suitably reflective conditions, we would ascribe – the relevant kinds of liability-responsibility to ourselves, and in that way take on the attendant liabilities, either positive or negative. Pamela Hieronymi is one author who defends this kind of picture about the bases of responsibility for belief. She argues that person A can be held responsible for involuntarily acquired attitudes, if those attitudes embody what A would, upon reflection, affirm as her own 'take on the world'. The person inculcated into a sexist belief-system is liable to blame, on Hieronymi's view, if and insofar as he would reflectively affirm the morally defective parts of that belief-system.<sup>27</sup> Hieronymi thus rejects the notion that liabilities which may be associated with beliefs are only properly ascribed if the agent bears causal responsibility for her beliefs. There is a noteworthy rationale for following Hieronymi, in rejecting this connection, which has been explored at length in Tony Honore's work on liability in tort law. Honore's view, in essence, is that it is in our interests to be held responsible for our conduct, since the long-run benefit of being praised and rewarded when our actions play out favourably overrides the immediate cost of being blamed or punished when they do not. The form of life and society that is possible when we can be credited and rewarded for achievement is, for each of us individually – and also, crucially, for all of us together – worth the cost of our being held to account for our failures in conduct. And this is true even though specific cases of failed conduct (e.g. someone's accidental but negligent causing of a road accident) are not within the control of the person we hold accountable.<sup>28</sup> A parallel rationale can be used to characterise the relationship between beliefs and liabilities. When a person's beliefs are laudable, he will naturally feel that he should be credited. After all, the fact that his beliefs come to him unbidden does not make them less true, or rational, or discerning, or subtle – and for these or any reasons laudable. In short, we have an apprehensible basis *qua* self-interested agents for wanting to be held responsible for our beliefs.<sup>29</sup> There is more at stake here, however, than the credit we might each desire for ourselves. Many of the social incentives we have to be

---

<sup>27</sup> Pamela Hieronymi, "Responsibility for believing", *Synthese* 161 (2008): 357-73; for another account along broadly similar lines see Angela M. Smith, "Responsibility for attitudes: activity and passivity in mental life", *Ethics* 115 (2005): 236-71.

<sup>28</sup> This is a broad-strokes summary of the central argument which is advanced in Tony Honore, "Responsibility and luck: The moral basis of strict liability" in *Responsibility and Fault* (Oxford: Hart Publishing, 1999).

<sup>29</sup> A similarly-oriented discussion about the general desirability of being counted as responsible for one's conduct is presented in John Gardner, "The mark of responsibility", *Oxford Journal of Legal Studies* 23 (2003): 157-71.

careful, inquisitive, non-credulous, accurate in our recall, precise in our comprehension, and epistemically virtuous in these and other respects, would evaporate, with lamentable consequences, if people were not as a matter of course credited for having true (or rationally acquired, or well-founded) beliefs. Which is to say, there is a form of *intellectual* life and society that becomes possible when are credited and rewarded for doxastic success; and again, for each of us individually, and for all of us together, this will be worth the unwelcome costs of our being held to account for our doxastic shortcomings, even though those shortcomings in general originate beyond the compass of our agential control.

Taking responsibility for our beliefs is thus, so we might say, a matter of taking the bad with the good. Of course, the individual will not typically regard her own beliefs as bad or blameworthy. If I judge that believing *p* would be irrational or credulous, or if I think that *p* is the kind of belief that leads people towards actions that are (by my own lights) egregiously wrong, then presumably these judgements make *p* (for me) ineligible *qua* belief. But taking responsibility for one's mental life doesn't mean thinking about individual belief-tokens, and asking whether one may be blameworthy for believing *p*<sub>1</sub>, *p*<sub>2</sub>, *p*<sub>3</sub>, etc. It means consciously adopting a proprietorial stance in relation to one's own thinking and believing, despite the possibility of unwelcome consequences being incurred in light of such a stance. We may be held responsible for our beliefs, even though we are not agentially responsible for acquiring them, because from a duly reflective point of view, mindful of the overriding goods that are at stake for ourselves and others, we would allow ourselves to be answerable or accountable in that manner.

One might worry that there is an unsustainable tension in this position. It may appear that we are expecting individuals to adhere to a mistaken conception of their own agential nature; to falsely believe that they *do* in fact bear agent-causal responsibility for their beliefs, such that they are liable for those beliefs in turn. At the very least, the critic might want us to concede that an account like this relies upon some amount of illusion. If person A allows herself to be liable to blame or penalty vis-à-vis her beliefs, she cannot at the same time, and in the same cast of mind, consciously attend to the thought that she, as the one who stands to suffer, is at bottom *innocent* in relation to the thing that is blameworthy or liable to penalty. As Saul Smilansky argues (not just about belief, but about taking responsibility for ourselves in general), we will not be able to sustain practices that involve valuing people and their behaviour, or feeling due remorse, or aiming for integrity, if we are constrained by a sense of ourselves as things that are merely acted upon. In order to sustain such practices we must recognise that it is useful for all (or most) of us to have an illusory sense of our own agential powers.<sup>30</sup> These are reasonable concerns, as far as it goes, but it

---

<sup>30</sup> See in particular chapters 7 and 8 of Saul Smilansky, *Free Will and Illusion* (Oxford: Clarendon Press, 2000).

is important to be clear about which parts of our account depend upon illusion. The Honoré-style argument is intended to deny the necessity of the justificatory link between agent-causal responsibility and the ascription of blame or liability. We count ourselves accountable for believing because otherwise there will be an unfair asymmetry. We reap the undeserved benefits when we (non-volitionally) do well in our doxastic lives, so we should accept the costs when we (non-volitionally) do poorly in this respect. The moral justification for our liability-ascription schema does not rely upon participants living under illusions about the non-volitionality of belief, even though such illusions may be needed to sustain the social economy of motivation and expectation that such a schema requires in order to function.

## 8. Free speech after involuntarism

In §6 I argued that we are not responsible *capacity-wise* for our beliefs. And in §7 I argued that we may nevertheless be held responsible, *liability-wise*, for our doxastic lives. So where does this leave free speech theory? In this final section I will expand on my claims about (i) how involuntarism bears on the free speech arguments that I outlined in §3, and then (ii) how my re-conceptualisation of the idea that we have minds of our own allows it to do the work in free speech theory that I discussed in §§4–5.

(i) Our lack of capacity-responsibility with respect to our beliefs undermines the *prima facie* appealing arguments for free speech that we find in the work of Mill, Scanlon, and Dworkin. We do not exercise agential control over our doxastic lives. We should therefore, by default, adopt a sceptical stance towards Mill's argument about the value in our intellectual flourishing, and towards the Scanlon/ Dworkin argument about government legitimacy. Why? Because in order to formulate putative justifications for free speech, these arguments appeal to an implausible thesis about human capacities. Should we conclude, in addition, that these authors were misguided in their views about human nature? Maybe, although we needn't leap to that judgement on the basis of what I've said here. The underlying worries which propel the arguments from Mill, Scanlon, and Dworkin, are about the moral wrongness of (certain species of) relationships involving domination and control: conformist tyrannies of the majority (for Mill), and neurotically censorious, bureaucratic nanny-states (for Scanlon/Dworkin). In their attempts to formally characterise the wrongness of such relationships, the authors invoke a familiar and *prima facie* agreeable thesis about our intellectual agency. This is where things go awry. The thesis they appeal to is a makeshift approximation of a certain part of our uncritical self-image. And when it is scrutinised, that part of our

self-image – the part where we see ourselves as the authors of our doxastic and mental lives – reveals itself to be a much more partial and fragmented thing than we (most of us) initially and uncritically supposed. But the proponents of these arguments can grant all this while retaining the moral misgivings (concerning relationships of domination) out of which their arguments were born.

For those who share Mill's perspective and his concerns, the moral worries about oppressive conformity and tyrannies of the majority can be fleshed out, without any need for bold claims about intellectual agency, just by stressing the utilitarian costs of these things. Those who think of themselves as intellectually autonomous free-thinkers may be labouring under misapprehensions about how and from whence their thoughts and beliefs originate. But it may still be true, for all that, that the kind of self-image they possess *is* instrumental in bringing forth the innovation and eccentricity in thought that Mill prizes so highly. And moreover, it may still be true that the experience of conformity breeds genuine suffering in people who (wrongly) conceive of themselves as free-thinking individuals. In short, both the immediate and the long-term utilitarian desiderata that animate Mill's argument might indeed be at stake in the question of whether societies acquiesce to conformist majorities, or try to counteract them via free speech policies. The notion that we have minds of our own is adjunct to, but not necessarily inextricable from, the moral considerations that underwrite Mill's argument. Proponents of the Scanlon/Dworkin argument, on the other hand, can pull back from the hard-line version of the claim about conditions of legitimate governance that I've attributed to them. The objection to the censorious government, recall from §3, was that subjects who regard themselves as having minds of their own (in the sense of being independent, rational, and autonomous) cannot reasonably consent to living under that government's rule. However, so the proponent of the argument can say, this condition of government legitimacy may still apply even if it is in fact the case that subjects are mistaken in understanding themselves as having minds of their own. One may want to say, for instance, that legislation which violates this condition assails its subjects' dignity not because it rides roughshod over *their* capacities, but because it is predicated upon a perniciously asymmetric (and false) notion: that the governing parties *do* possess some capacity for independent and autonomous judgement which the governed parties lack. Again, the notion that we have minds of our own is adjunct to, but possibly extricable from, the normative core of the argument.

I've not said nearly enough to substantiate any confidence in the viability of these arguments in their reconstituted forms. Nor have I given any reason to subscribe to the other social good arguments or individualistic arguments for free speech mentioned in §2. My view, though I've not argued in its defence, is that there *is* plenty to be said on behalf of free speech. In order to show what there is to be said on free

speech's behalf, we have to address several sets of questions, e.g. about (i) how the case for free speech is properly elaborated, and (ii) what activities should be recognised as 'speech' for the purposes of free speech, and (iii) what it would require for free speech to receive its due in the *de jure* policies and *de facto* politics of contemporary societies. However, since I won't be taking up any of these questions here, I will express my (first) conclusion in a conditional form. There are some theorists who are sceptical about the mainstream liberal view, according to which free speech is an important ideal/desiderata in a just political order. Some argue that the common conception of free speech as an outworking of *de jure* government neutrality is paradoxical. Free speech, they say, is only ever an *ad hoc* family of measures aimed at resisting grossly oppressive forms of government.<sup>31</sup> Others argue, in a more iconoclastic spirit, that free speech is nothing but a shibboleth, opportunistically invoked when it is rhetorically expedient in the promotion of some partisan agenda.<sup>32</sup> My claim is that those who see free speech as something more than this – as a *bona fide* principle or right that governments are duty-bound to honour – must articulate their bases for this view without appealing to that familiar and attractive thesis, about our intellectual agency, that the Mill and Scanlon/Dworkin arguments (in their most straightforward guises) are reliant upon.

(ii) If that challenge can be met, we return to the situation that I described in §4, where we have granted that free speech is an important value, and where our task is to determine how it should be balanced with other values. My claim in §5 was that the idea that people have minds of their own shapes and organises the responsibility-ascription framework involved in these determinations in two ways. Firstly, in cases where speech leads to harm through an ostensibly linear sequence of events, this idea helps to guide the rendering of specific verdicts about which participants are accountable (and to what extent) for particular harms. Secondly, this idea can be used to defend the whole enterprise of attempting, in the first place, to distinguish between the cases that are more like persuasion and those that are more like provocation or incitement. It is, so I've said, precisely *because* we understand people as having minds of their own, rather than being completely pliable in response to the speech of others, that such distinctions are needed. In §7 I explained why we *can* hold each other responsible for what we think and believe, even though we are *not* ultimately responsible capacity-wise for our mental lives; in short, because we would (or should) ascribe such responsibility to ourselves. And as I said earlier, this ascriptive reformulation of the idea that we

---

<sup>31</sup> See for instance Chapter 8, "The paradoxes of liberalism and the failure of theories justifying a right of freedom of expression", in Larry Alexander, *Is There a Right to Freedom of Expression?* (New York, Cambridge University Press, 2005).

<sup>32</sup> Most notably Stanley Fish; see especially the title essay in *There's No Such Thing as Free Speech, and it's a Good Thing Too* (Oxford: Oxford University Press, 1994).

have minds of our own should be sufficient to vindicate the guiding role that this idea plays (as explained in §5) in the attribution of responsibility for harm in a free speech policy schedule.

What remains to be made explicit, even if one accepts all this, is how we make the move from the claim that ‘we *can* hold people responsible (in free speech policy) for what they believe’ to the stronger claim that ‘we *should* hold people responsible (in free speech policy) for what they believe’. Consider the case of pornography again. Just because we now have an account which tells us that it’s justifiable to hold consumers of pornography responsible for pernicious effects that pornography has upon their attitudes, and for the harmful consequences that ensue, it needn’t automatically follow that, all things considered, we *should* allow our responsibility-ascriptions to be determined via this account. After all, one upshot of doing this is to suggest a limitation on the extent to which the producers and distributors of pornography can be made answerable for the harms at issue, even though these harms are, in at least some cases, quite ascertainably likely to result from the influence pornography exerts on its consumers. And why should we accept this? Why not do something bolder in response to the finding that people’s intellectual agency is meagre and indirect? If people are indeed hapless products of the belief-forming environments they find themselves in, why not conclude that full and unqualified liability *should* redound to those who influence these environments in a way that increases the chances of harm eventuating?

The problem with this line of thought, in essence, is that it pulls us into a vicious regress, indeed, a regress whose viciousness owes to the very fact that it makes it *more* difficult to further the aim of harm-prevention or -deterrence through our ascriptions of responsibility. Where A is influenced by B towards beliefs or attitudes that later give rise to harmful acts, there is no non-arbitrary way to fix in place the verdict that accountability should fully redound to B. Why? Because B will have been influenced in turn by C, who will have been influenced in turn by D (and so on), and if A, B, C, D, etc. are *all* adjudged accountable for the harms at issue, then no-one can be usefully singled-out for accountability in the name of harm-deterrence or -prevention. And the aim of harm-deterrence and/or -prevention is what drives the whole inquiry into responsibility-ascription in the first place. What we require, in seeking to ascribe responsibility for belief, is an ascriptive framework which contains the resources for discriminating between more and less pernicious modes of communicative influence – between, say, incitement and persuasion. That requirement is what gives us a positive reason to sustain our practices of holding people responsible, in free speech policy, for how they think and believe in response to the influence of others. What involuntarism calls into doubt is whether we can *justifiably* sustain those practices. This is where my revisionist account of responsibility-ascription in §7 is meant to do its work. On this account, we can



justifiably hold people accountable for their mental lives, even though they are not responsible capacity-wise for their mental lives, because all of us would (or should) take on this kind of accountability.

## **9. Conclusion**

My aim here has been to explain the relationship between intellectual agency and responsibility for belief in free speech theory. The idea of human beings as intellectually self-governing creatures has often been invoked as part of the justification for free speech, through arguments like those we find in Mill, Scanlon, and Dworkin. What has been less well-recognised is the role that this idea has to play in translating free speech into legal policy, by guiding a framework for ascribing responsibility for certain kinds of harms causally associated with speech. The crucial question, so I've suggested, is what becomes of these relationships once we recognise that the initial view about our intellectual agency is, in its characteristic form, untenable. I have argued that this will undermine the justificatory force of claims about our intellectual agency in relation to free speech. However, I have also explained how we can reformulate an account of our intellectual agency, in which we may still be held responsible for what we believe, even though we do not exercise agential capacities as doxastic subjects. This ascriptive account of intellectual responsibility still has a role to play in guiding the legislative outworkings of free speech theory. When communicative interactions causally contribute to harmful outcomes we have to formulate policies for assigning blame and liability. One crucial thing these policies will have to do is ascertain the different modes of communicative influence – incitement, provocation, persuasion, suggestion, etc. – that may be in operation in different cases. So far as we are concerned with specifying the fine-grained detail of such policy, my account of how we can be held responsible for our beliefs is only one thing to be considered among others. However, my account provides an essential reference point for when we are trying to defend and make sense of the underlying ascriptive framework that such policy presupposes.