

Three questions for liberals

Richard Pettigrew
Richard.Pettigrew@bristol.ac.uk

January 3, 2024

Abstract

In this paper, I ask three questions of the liberal. In each, I fill in philosophical detail around a certain sort of complaint raised in current public debates about their position. In the first, I probe the limits of the liberal's tolerance for civil disobedience; in the second, I ask how the liberal can adjudicate the most divisive moral disputes of the age; and, in the third, I suggest the liberal faces a problem when there is substantial disagreement about the boundaries of the rational and the reasonable.

In November 2023, Cass Sunstein wrote about liberalism for *The New York Times*, and then expanded on his piece in an academic paper (Sunstein, 2023a,b). Both essays take the form of a list: thirty-four propositions that the liberal professes. No doubt many friends and foes of liberalism will question the boundaries Sunstein draws for the position, and the sheer scope of the topic—to circumscribe a political approach discussed and practised in myriad different ways over a period of centuries—means that he cannot adjudicate every border dispute. But one distinctive feature of his vision of liberalism is how inclusive it is: as he says himself, Sunstein includes James Madison, Martin Luther King Jr., Margaret Thatcher, and Ronald Reagan on the political side, and Mary Wollstonecraft, Friedrich Hayek, Martha Nussbaum, and Robert Nozick on the philosophical side. Nonetheless, there are many who still stand outside this circle, and one of Sunstein's early claims is this:

#8: Liberals are aware that, all over the globe, liberalism is under siege.

This claim is also the target of the present issue of this journal. My aim in my contribution is to say something about why it might be true, if indeed it is. Yet I am no historian nor political scientist, and so my approach is different from theirs. I want to ask three questions of the liberal. In each, I try to fill in philosophical detail around a certain sort of complaint raised in current public debates about their position. I do not know the extent to

which these complaints play any causal role in the current march of illiberalism, nor even whether they play a role at all. I suspect they played no role in instigating it, but they might well play some small role in sustaining it and making it seem palatable to those who would typically oppose it. And if that is indeed a possibility, it is important to consider whether and how they might be answered. Throughout, I will use Sunstein's lists to identify the tenet of liberalism that is under siege in these complaints.

1 Some preliminaries: theories of rational choice

I begin in a rather surprising place: the theory of rational choice. At some point in each of the three questions I wish to raise, we will have cause to use some of the concepts and terminology of rational choice theory: prospects, states, and outcomes; preferences, probabilities, utilities, and attitudes to risk. These are terms of art, but they are extremely useful to have to hand, for they allow us to be much more precise when we discuss what individuals want, how they would and should choose on the basis of what they want, and so on. And such discussions are central to liberalism.

So let me illustrate these concepts by giving an everyday example of a choice someone might make, and labelling its different components with these terms of art, just as an anatomy textbook might label different parts of the ear.

Suppose I offer you a bet that pays out £10 if it rains in Bristol tomorrow and nothing if it doesn't. And suppose I'll charge you £7 to enter into this bet with me. Then you face a choice: (i) accept the bet; or (ii) reject it. We call these two options *prospects*. And how a particular prospect turns out for you depends on which of two situations you're in: (a) one in which it rains tomorrow; (b) the other in which it doesn't. We call these situations *states of the world* or simply *states*. In the state of the world in which it rains, accepting the bet gains you £3 overall, you'll be £7 down from the entry fee, while in the state in which it doesn't rain, it loses you £7. We call these the *outcomes* of that prospect at those states. In both states rejecting the bet leads to the same outcome, namely, the one in which you gain nothing and you lose nothing. A prospect and a state of the world together fix an outcome.

How should you choose between the prospects? You should pick the one you prefer, or pick either if you prefer neither. Your *preferences* rank the prospects. But how are you to set your preferences? The standard story, which comes down from Daniel Bernoulli (1738 [1954]), says that they should be fixed by (i) how much you want or desire or value the different possible outcomes and (ii) how strongly you believe in the different possible states of the world on the supposition that a given prospect is chosen. Whether you prefer accepting to rejecting the bet or prefer rejecting

to accepting or are indifferent between them depends on how much you want to gain the £7, how much you want to avoid losing the £3, how likely you think it is to rain, and so on. We represent your desires, wants, and values by your *utility function*, which assigns a number to each outcome that measures how strongly you want it. So, since money has diminishing marginal utility, you might assign a utility of -10 to the outcome in which you lose £7, a utility of 2 to the outcome in which you gain £3, and a utility of 0 to the outcome in which you neither lose or gain. We represent your beliefs, on the other hand, by your *probability function*, which assigns, for each prospect, a probability that measures how likely you take each state of the world to be on the supposition that the prospect is chosen. So you might think it's 20% likely to rain tomorrow on the supposition you accept the bet, and 80% likely to stay dry on the same supposition; and you might give the same probabilities on the supposition you reject the bet because, in this case, the state of the world is independent of which prospect is chosen—whether you bet or not doesn't affect whether it rains or not.

The standard story says that you should prefer one prospect to an other if the first has higher expected utility than the second relative to your probabilities and your utilities. The *expected utility* of a prospect is obtained by working through each state of the world one at a time, taking the utility of the prospect's outcome at that state, weighting it by the probability of that state on the supposition that the prospect is chosen, and then summing up these weighted utilities. So, for instance, given the utilities and probabilities just given, the prospect of accepting the bet has expected utility of $(0.2 \times 2) + (0.8 \times -10)$

That's the standard story, but there are alternatives around. Some of them hold that your preferences are fixed not only by your probabilities and utilities, but also by your attitudes to risk.¹ The details don't matter here. All that needs to be said is that these alternatives tell you to rank a prospect by a quantity that is determined not only by the probabilities you assign to states of the world on the supposition of that prospect and the utility you assign to the outcomes, but also by your *attitudes to risk*. The expected utility of a prospect is the sum of the probability-weighted utilities of the outcomes. Let's say that this is the *risk-neutral* approach. A *risk-averse* approach would give more weight to the outcomes with the lowest utility (the worst-case outcomes) and less to the outcomes with the highest utility (the best-case ones) than the neutral approach would give; and a *risk-inclined* approach would give more weight to the outcomes with the highest utility and less to the outcomes with the lowest. As a result, a risk-averse person might prefer a less risky option, which is much more likely to give a middling utility, over a more risky option, which has a good

¹For some risk-sensitive decision theories, see (Quiggin, 1982; Chew, 1983, 1989; Quiggin, 1993; Buchak, 2013; Bottomley & Williamson, ta).

chance of giving a high utility but also a good chance of giving a low utility, while a risk-inclined person with the same probabilities might prefer the riskier option.

With all this terminology in place, we can now begin to formulate our questions.

2 The first question: liberal disobedience

The first principle in Sunstein's list of liberalism's tenets is this:

#1: Liberals believe in six things: freedom, human rights, pluralism, security, the rule of law and democracy. In fact they believe in deliberative democracy: an approach that combines a commitment to reason-giving in the public sphere with a commitment to accountability.

It has become an online joke raised to the level of a meme that liberals will tell you the only way to change society is to vote. But of course that isn't true. Liberals also support: mobilizing citizens to lobby for a cause or to draw attention to it or to formulate a policy concerning it; bringing lawsuits against the government to challenge past misdemeanours, deter future ones, and place on the official record a verdict about their conduct; uncovering these misdemeanours and publicizing them through investigative journalism; and so on. But what does the liberal say if a society is liberal in many respects, but also illiberal in some, and the strategies just described are insufficient to shepherd it the final distance towards the fully liberal society they'd like to see? Will they then endorse actions beyond those just described? And, if so, to what extent are they prepared to violate in the short-term the six pillars of liberalism that Sunstein lists—freedom, human rights, pluralism, security, the rule of law, and democracy—to secure a liberal future that enshrines those tenets in the long-term?

Stated like this, this is the question of civil disobedience, its scope and its permissibility. And liberals have had plenty to say about it. Candice Delmas's recent book, *A Duty to Resist*, is perhaps the state of the art from the liberal perspective (Delmas, 2018). She characterizes the prevailing Rawlsian (1999 [1975]) account as follows:

[C]ivil disobedience is a conscientious, public, nonviolent breach of law undertaken to persuade the majority to change a law or policy in a nearly just society. Rawls understood publicity to require that agents give authorities fair notice of their planned disobedient activity, act in public, and appeal to the community's shared conception of justice. He thought that nonviolence excluded the use of force (or coercion, in Rawls's usage) and actual or likely infliction of harm against persons. In addition,

agents of civil disobedience are supposed to accept, and even seek out, the legal consequences of their actions. In doing so they demonstrate their general “fidelity to law,” endorsement of the system’s legitimacy, and belief that the latter generates a moral duty to obey the law. (Delmas, 2018, 25)

Delmas finds this account unduly restrictive, both in the kinds of disobedience Rawls permits and the circumstances in which he permits it. She replaces it with her own, which holds that both civil and uncivil disobedience—what she calls together *principled disobedience*—is sometimes justified in a nearly liberal society; and indeed, she later argues that, in some cases, it is not only permissible, but obligatory.

There is a place for uncivil disobedience in liberal democratic societies when the following conditions apply: the public is assured of the state’s commitment to respecting everyone’s full and equal status, a commitment typically embedded in a constitution or other basic law that guides institutional design and lawmaking; some citizens are effectively (*de facto* but not *de jure*) denied full and equal status; and the injustice of this denial is not publicly recognized, perhaps because that injustice is not deliberate but results from the interplay of social practices and institutional structures, as in cases of structural injustice. (Delmas, 2018, 64)

On both of these accounts, principled disobedience is aimed primarily at educating the public, trusting that doing this will be sufficient to bring about societal change. It draws citizens’ attention to an injustice present in their society, which, it is assumed, they will then vote to overturn through the democratic process. Rawls explicitly talks of the persuasive aim of civil disobedience, and his examples are drawn from the civil rights and anti-war movement protests of the mid-twentieth century; Delmas also requires that there is an injustice that is not fully recognised by the majority, and her examples include Daniel Ellsberg’s Pentagon Papers, Edward Snowden’s whistleblowing, protests staged by Pussy Riot and Femen, and the riots precipitated by police brutality in the banlieues of Paris and the cities of the United States after Théo L.’s rape and George Floyd’s murder, both at the hands of law enforcement officers.

These accounts assume that, while the society commits the injustice that the disobedience targets, and so for this reason is illiberal, the mechanisms by which the disobedience might lead first to majority awareness of the injustice and then from that to democratically chosen change are intact in the way the liberal would like them to be. Ellsberg and Snowden intended that they teach US citizens of their government’s malfeasance, and the citizens then vote them out; Pussy Riot intend that we become aware of the patriarchal nature of our society, and choose collectively to change it through the

ballot box. That is, the sorts of principled disobedience that both Rawls and Delmas discuss rely for their efficacy on a functioning deliberative democracy. But this raises the question: what will the liberal permit if the aspect of liberalism that has not been fully achieved in the society is precisely such deliberative democracy? What sort of principled disobedience has any chance of working in such a case, and can the liberal countenance it?

To be concrete, let us consider a society much like many of those we call liberal democracies today. In it, there are weak restrictions on political campaign financing, allowing the very wealthy to influence which political candidates get a chance to run and what their platform will be, and it is possible and permissible to pay for greater personal access to politicians; there is a revolving door between government, the civil service, and other civic institutions, on the one hand, and the private sector and news media industries, on the other, creating incentives for members of each side to fall in line with the policies of the other, since that will best serve their material needs; there are few rules governing when the media is able to put questions to politicians, thereby making much-valued access to the key players contingent on not posing questions that are too challenging; the conduits for sharing knowledge, such as news media and repositories of collective knowledge, and the fora for hosting public debate, such as social media sites, are owned by the extremely wealthy. Such a society, I think, would not live up to the liberal ideal that Sunstein describes as deliberative democracy.

In such a society, any disobedience of the sort that Rawls and Delmas countenance, which aims only to highlight these illiberal features, will be ineffective because the mechanisms that would disseminate the information it hopes to highlight aren't functioning well. The media has an incentive to selectively report any protests, to place excessive emphasis on any episodes of disorder that occur during them, and to launch ad hominem attacks against the protesters, based perhaps on ugly episodes from their past or the lives of their associates, deflecting attention from the claims of the protests. The norms and functioning of the social media sites that host the public discussion are under the control of owners with an incentive to maintain their position of power and influence; and there is very little transparency about how they do function, since their algorithms and other processes are protected as intellectual property. And the integrity of the politicians is compromised, owing favours to the wealthy who helped their election campaign and hopeful that a well-remunerated role in the corporate world awaits them after their time in politics.

So a different form of principled disobedience is required to effect this sort of change. What might it be? I have no roadmap. Perhaps it would involve seizing control of the online public spaces and the conduits of information, not briefly, as hacker activists such as Anonymous have sometimes done, but for the long term, with a view to building a more democratic, de-

centralized ownership structure that better serves a deliberative democracy with a large citizenship.

What should the liberal think of such an action, if that really is the best hope for bringing about a truly liberal state? Delmas argues that, as well as the sort of civil disobedience that Rawls condones, the uncivil disobedience that she permits can survive the standard objections against it. One such objection is that, in a near-just state, citizens are morally obliged to obey the law. And one standard argument for this is that those who break the law are free riders. According to this argument, citizens in a near-just state are participants in a mutually beneficial cooperative scheme: they receive the benefits of personal safety and security, freedom to pursue their own interests, and so on, but these benefits are only possible if a sufficient proportion obey the law. In this situation, if you don't then obey the law, you make an exception of yourself in a way that violates requirements of fairness. We might assume that each person could benefit for themselves were they to break the law in some way against a background of others obeying it; and so, if you do break the law, you are doing something that others would like to be able to do, but cannot all do without everyone suffering. And this, it might be said, is immoral. However, as is often pointed out, making an exception of yourself in this way is only immoral if you break the law in order to benefit yourself alone or yourself and some small class of associates. If nearly everyone stands to gain from your law-breaking, it no longer constitutes an immoral form of free-riding. And a group that seized control of media infrastructure in the way I described above could claim that everyone does stand to benefit from their law-breaking.

A graver concern about the sort of law-breaking that is involved in uncivil disobedience, in particular, is that it can threaten the stability of society. Indeed, many of the conditions that Rawls imposes on acceptable civil disobedience, such as informing authorities in advance, acting in public, and seeking out legal redress for one's actions, are intended to maintain such stability by signalling that the disobedients respect the rule of law. But these conditions are often not met in the sort of uncivil disobedience Delmas condones, and certainly not in the sort of law-breaking I described above, which might end up being our only route from where we are now to where the liberal would like us to be. It is not unreasonable to think that the risk of instability increases with the severity of the lawbreaking, and seizing control of very valuable digital assets is surely very severe.

One problem for the proponent of such principled disobedience is that it is an empirical question which actions raise the probability of instability, and how by much. And presumably it is a question over which reasonable people with roughly the same evidence might disagree, since the evidence is messy and pulls in various directions and there is no single way in which it must be weighted and aggregated. And, in the case we're considering, in which digital infrastructure is seized by activists with a view to placing it

under more democratic governance, it is not just the risk of instability we run, but also the risk that the infrastructure fall into more illiberal hands, rather than more democratic ones—such is not an unknown consequence of revolutions that seize control of important assets. But if there is such widespread disagreement about the wisdom of such an action, and the liberal’s favoured methods for resolving such disagreement, namely, deliberative democracy, is precisely what is absent, it’s unclear how the liberal could say that such action is legitimate.

What’s more, even among those who agree on the probability that instability or illiberal control will result from the disobedience, there might still be disagreement about whether to condone such actions, since they might have different attitudes to risk, so that the risk-averse prefer not to resist in the way outlined, accepting the less-than-fully-liberal situation in order to avoid any increase in the risk of instability, while the more risk-inclined prefer to resist, reasoning that, on balance, the increased risk of instability is worth it in order to make available the possibility of the fully liberal society. And so again there is widespread disagreement about the wisdom of the action, and the liberal’s democratic methods for resolving such pluralism into decisive action aren’t fully available.

So the first question to the liberal is this: When the only principled disobedience that stands a chance of leading to a fully liberal state must go beyond the sort of protest or whistleblowing that serves to highlight an existing injustice because the mechanisms for disseminating information publicly are compromised, what level of disobedience is permissible and how should we handle disagreement between citizens about the risk of instability or despotism they are prepared to run to achieve a fully liberal state?

3 The second question: the anarchy of public reason

In the fourth of his claims, Sunstein quotes Lincoln’s Peoria speech of 1854, and picks out one line for special attention:

#4: Liberals especially like this: “No man is good enough to govern another man, without that other’s consent.”

It has been the work of the so-called public reason strand of liberalism in the twentieth and twenty-first centuries to make clear what this means.² Here is not the place to describe the many views that have emerged from this project. Instead, I wish to articulate a general concern about liberalism that has come to the fore recently, and that is best understood as an objection to some public reason approaches: certain disagreements in society

²For an overview, see Jonathan Quong’s (2023) entry in the *Stanford Encyclopedia of Philosophy*.

are simply too great to be adjudicated by the sort of processes that public reason liberals endorse.

Must we say, as Lincoln seems to suggest, that each person subject to a government's power must *actually* consent to being governed by it? Surely this is too strong, and no government will be legitimate if we require this for its legitimacy. Perhaps then we require only that they *would* consent to it, *if asked*? And yet, for any system of government or body of legislation, many people realise that there are alternative systems or bodies under which they, with their particular talents and desires and resources, would fare better; and so they will not consent to the present system if these others are available.

Two roads lie ahead for the public reason liberal at this point. They may follow John Rawls (1999 [1975]) and require that each person would consent to the system from behind a veil of ignorance, where they are stripped of all knowledge about the position in society they will occupy. The problem with this is that, behind that veil, they find themselves facing a choice while in a state of uncertainty, and so they require a principle, or decision rule, by which to make a choice in such a state. Rawls, famously, opts for the so-called Maximin decision rule (Rawls, 1999 [1975], §§26-28). This rule is maximally risk-averse in the sense described above, for it says you should look only at how a prospect fares in its worst-case scenario, and you should choose a prospect whose worst-case scenario is best. So, for instance, if you're offered £5 for sure or a gamble that gives you £4 if it rains tomorrow and £1,000,000 if it doesn't, you should choose the sure £5 since it gives you £5 in its worst-case scenario, while the gamble gives you only £4 in its worst-case scenario. But there are many alternative decision rules for situations of uncertainty, and Maximin is not in any sense required by considerations of rationality. Once again, then, there is no reason to think that there is any system of government to which each person will consent from behind the veil of ignorance: those who use different decisions rules will favour different systems.

Alternatively, the public reason liberal may follow Gerald Gaus instead of Rawls. Where Lincoln says that it is necessary that a person consent to be governed, Gaus (2010) spells this out as saying that a system of government is legitimate if each reasonable person subject to this system prefers it to no system at all—that is, they prefer being governed in this way to living in a state of anarchy.

There are at least two sorts of worry about this version of public reason liberalism:

(1) There are surely illiberal systems of government that persecute me in various ways that I would nonetheless prefer to no system of government at all. Provided the discrimination against me and the ways in which I am arbitrarily treated are reasonably contained and do not involve very serious harm, I might well take my chances with that rather than throw my lot in

with an ungoverned state of nature. In other words, Gaus sets the bar very low; many systems of government that the liberal disavows will likely clear it.

Of course, Gaus will respond, the bar must be set so low so that some system of government will clear it; we cannot require that everyone must be positively happy with the system of government, nor that they condone all of its policies, since there will be none that does that for everyone. The most we can hope for is that we find one that all agree is better than nothing at all. Now, it may well be true that public reason liberalism can only be saved by setting the bar so low. But the illiberal consequences of doing that thereby tell against public reason liberalism as a project, and suggest that this way of making precise what Lincoln demanded at Peoria is not a route to true liberalism.

(2) Even lowering the bar so far provides no guarantee that we can find a system of government to which all consent in Gaus's sense. There are three ways this might happen that I think lie behind some contemporary concerns about liberalism.

(2a) We can easily imagine a moral dispute over a particular practice with the following feature: partisans of one side say that *permitting* the practice is morally abhorrent, while partisans of the other side say that *forbidding* it is morally abhorrent. Historical examples of this might involve practices of religious worship: I think it is an abomination in the face of God to worship in this way, while you think it is an abomination in the face of God not to worship in this way. But there are other sorts of disputes that have this structure that are more prominent in contemporary political debate: disputes over the right to abortion and bodily autonomy for someone who is pregnant; disputes over the rights of trans people, and particularly trans children, to gender-affirming healthcare, institutional recognition, and bodily safety; disputes over the use of culturally or religiously significant land. Some who would ban abortion even in cases of rape and severe risk of harm to the pregnant person liken the practice to murder and consider it just as morally abhorrent; some who would permit it in such cases argue that a prohibition places people at substantial and unnecessary risk of harm and death and consider that morally abhorrent. The problem is that, if the moral abhorrence on each side is severe enough, and if the practice, if permitted, would be sufficiently widespread, then each side will find the other's favoured legislation intolerable, and perhaps intolerable to such a degree that anarchy would be preferable.

Now it's easy to see why they'd find it intolerable, but why think they might prefer anarchy? Well, it is important to remember that, if the practice is prohibited, then the state's power and monopoly on coercive violence will stand behind that prohibition; and, if that's the case, it will be near to impossible for a private individual to undertake the practice. On the other hand, if the practice is permitted, then the state's power and its

monopoly on coercive violence will stand behind that permission; and, if that's the case, any attempt to curb the practice by those who find it abhorrent will be thwarted by the state; after all, each set of laws comes with the higher-order law saying it is impermissible to forcibly prevent someone from acting in a way permitted by these laws. So someone who favours the practice might reasonably conclude they stand a better chance of being able to practice it in a state of anarchy than in a governed state that prohibits it, while someone who abhors the practice might reasonably conclude they stand a better chance of being able to curb the practice in a state of anarchy than in a governed state that permits it.

Now, the public reason liberal might think that whatever force our partisans to this moral dispute would meet under anarchy would be worse than the force they'd meet under a liberal state. But this is an empirical question. It is one about which we might be quite uncertain; and, importantly for our purposes here, it is one about which different members of the public might significantly disagree.

(2b) So we will be unable to find a system of government that all prefer to anarchy if there is a moral dispute so severe that one side prefers anarchy to permitting the disputed practice while the other prefers anarchy to prohibiting it; and so, in this situation, there is no system of government to which all subjects consent. But I think such cases are rare. More common is the following situation: there is a series of moral disputes about a series of practices; those who morally abhor a particular practice from this series prefer anarchy to a state that permits that practice, while those who do not morally abhor it don't necessarily prefer anarchy to the state in which that practice is prohibited, but they do prefer anarchy to a state in which every one of the practices in this series is prohibited. For instance, the series of practices might include: eating meat and using non-human animal products for other purposes; conducting conversion therapy on queer and trans people; using the sacred sites of certain people for mining concessions; excessive use of fossil fuels to live at a particular level of luxury; discriminating against certain groups on the basis of religious beliefs; failing to take measures to prevent the spread of communicable disease; assisting the suicide of someone suffering from intolerable and untreatable pain from a terminal disease. For each of these practices, there might be members of a society who find it so morally abhorrent that they prefer anarchy to a system that permits it; and, while there is no-one who prefers anarchy to a system that prohibits any one of these, since they individually do not encroach too much on personal freedom, there are those who prefer anarchy to a system that prohibits all of them, since that system would be intolerably restrictive. And in such a case, there will be no system that is legitimate on Gaus's definition—there will be no system to which all subjects will consent, because there is no system that all prefer to anarchy.

(2c) Finally, it is a feature of all public reason accounts that the options

between which we must choose are systems of government or bodies of legislation, and not fully spelled out ways in which the society might develop under those systems or because of those laws. In the language of Section 1, they are what we called prospects; they are not what we called outcomes. That is, they do not determine the fully detailed outcome that will follow from them, but rather make different outcomes more or less likely.

How should we choose between prospects? As I explained in Section 1, the standard answer from the theory of rational choice is that we should choose a prospect with maximal expected utility. But, as I said there, some rational choice theorists deny this. Some say that you can also incorporate your attitudes to risk into your decision, so that a risk-averse person can give more weight to worst-case outcomes and less to best-case outcomes than expected utility theory requires, and a risk-inclined person can give more to best-case outcomes and less to worst-case ones.

Now, we might imagine a person with the following combination of attitudes to risk, utilities, and probabilities. In their attitudes to risk, they're risk-inclined. That is, the best-case outcomes loom larger in their decision-making than they do for the risk-neutral person, who ranks prospects by their expected utility, or the risk-averse person; and the worst-case outcomes loom smaller. As a result, they are more inclined to rank a more risky prospect, which opens the possibility of high utility but also the possibility of low utility, above a less risky one, which opens neither the possibility of high nor the possibility of low but only really the possibility of middling utility.

In their utilities, our imagined person cares primarily about their own well-being and that of their family, friends, and community. They might give some weight to the well-being of others, but it is much much lower than the weight they give to their close circle.

In their probabilities, this person recognises that their safety will be less robustly protected in an anarchic state than in a liberal democracy, but they think it likely they will nonetheless band together as a community to protect one another from external interference, which will mitigate this to some extent. And, what's more, they think the absence of a coercive state will allow them to build up wealth, power, and resources for themselves, their family, friends, and community in ways that a liberal democracy will prevent. After all, they reason that, while protections of private property provide safety for their own belongings and land, they also put the belongings and land of others beyond their grasp, and they are willing to risk the loss of their property to have the better opportunity to seize someone else's and thereby enrich themselves.

Combining these attitudes—the inclination to take risks, the limited circle of people whose well-being has substantial weight for them, and the beliefs about how they and their circle will manage in an anarchic structure—we obtain a preference for anarchy above all else. And that, of course, leads

to anarchy for the public reason liberal: if there is just one member of the public who ranks anarchy above all else, then there can be no system of government that every member of the public prefers to it.

And so our second question to the liberal is this: How do we make precise the demand that Lincoln made in Peoria in a way that ensures there is some system of government, and indeed a liberal one, to which each person governed by it consents?

4 The third question: disputed principles of reason

A common complaint, not about liberalism itself but about some recent governments in various parts of the world that claim to be guided by liberal principles, is that they are excessively paternalistic. They treat the preferences of certain groups—particularly those with less formal education, or with particular tastes or interests—as mistaken or corrupted in some way such that those preferences must be corrected or fixed up before the government takes them into account in its decision-making.

It was in this context that people complained when the British Prime Minister Gordon Brown called a Labour supporter “that bigoted woman” and the US presidential candidate Hillary Clinton placed half of Trump’s supporters in a “basket of deplorables”. And it is often said that liberalism lends itself to such attitudes, since the liberal’s tolerance for a plurality of approaches to life and systems of values is always tempered: *reasonable* approaches to life are tolerated and, so far as possible, accommodated within a liberal system; *unreasonable* approaches may be tolerated, but they are not accommodated. And yet, without a precise and neutral theory of what makes a preference reasonable, this exception to the liberal’s tolerance for difference can give them too much room to exclude.

So I anticipate that many would take issue with the tenet of liberalism that Sunstein’s states as his seventeenth principle:

#17: Most liberals are receptive to “nudges,” understood as freedom-preserving interventions such as warnings, reminders, and disclosure of information.

They would complain that nudges are inappropriately paternalist interventions on the part of a governing class who judge the preferences of less powerful citizens by a standard of reasonableness that is too deeply biased towards their own preferences, and does not encode an objective norm.

I think Sunstein himself has a natural response to this complaint. As he and his co-author Richard Thaler conceive of nudges, a nudger is a *means paternalist* but not an *ends paternalist* (Thaler & Sunstein, 2003, 2008). That is, they make no judgments about what ends you have. Rather, they try

to nudge you towards taking better means to your ends when certain psychological mechanisms, such as biases and bad reasoning, can make poorer means to your ends seem better to you than they really are.

Take an example: Suppose I have a bias in favour of more immediate gratification, and for that reason I will not contribute to the pension plan my employer offers, choosing instead to spend the money now. Judging that this is a poor means to the ends I actually have, the nudger presents the choice to me in a particular way. The option of contributing and the option of not contributing are both made fully available to me, so they have not limited my choice in any way. And yet they know that certain psychological mechanisms that are triggered by that particular way of presenting the options will likely lead me to choose to contribute to my pension, so that I end up taking what they hold to be better means to my ends. For instance, knowing that people are more likely to take the first option presented to them, absent a strong preference in favour of one of the other options, they place the option to contribute first on the form I must fill in (Carney & Banaji, 2012). So what the nudger judges unreasonable are means to ends, not ends themselves. As Sunstein and Thaler say, they are means, not ends paternalists.

Yet, spelled out more fully, this does not entirely escape the challenge. For, in order to determine how to nudge me, the nudger must take my preference for not contributing to the pension plan over contributing, discern from it my true ends, conclude that they are not best served by the means I prefer, and then present the choice in such a way that I am more likely to choose the better means instead. And to do this they need two theories of rational choice. The first is descriptive and allows them to take the preferences of someone, even if they are irrational, and extract from those preferences the different components that determine them: in particular, it must allow them to extract that person's ends (Thoma, 2023). So, in the case of the pension plan, they must be able to determine that, while I choose not to contribute, I nonetheless greatly value comfort in my old age. The second theory of rational choice they need is a prescriptive one and allows them to take the ends they've extracted in the first stage and determine the best means to take in pursuit of them—this is the sort of theory I described in Section 1. In the pension plan case, this will show that, by the lights of this prescriptive theory, and given the ends they've determined I have, the best means to my ends is to contribute.

But a problem arises if there is disagreement between the nudgers and the nudged about the correct prescriptive theory of means-ends rationality. Take the example of my pension contributions again. Why do I favour not contributing over contributing? Above, I said that I do it because I have a bias in favour of more immediate gratification, and the money I might contribute to the pension, which would secure for me gratification in the far future, could be used now to provide such gratification in the near future.

But is this really an unreasonable bias? After all, we often favour outcomes in which we get some good thing in the near future over alternative outcomes in which we get a little more of that good thing in the far future. If I had to choose now, I'd go for a single chocolate bar tomorrow over two bars in a year's time, and that doesn't seem unreasonable to me. In cases like these, economists say that I *discount the future*. For any amount of a given good—whether it's sums of money, intensity of a positive emotion, or bars of chocolate—I value it to a certain degree. Those degrees specify my ends, and they don't change over time. However, when I come to make a decision between different prospects, and the various possible outcomes of those prospects secure for me different amounts of each good at different times, I calculate how much I value a given outcome at that time by taking each moment in the future, looking at how much of each good I get at that moment in that outcome and the degree to which I value that, then apply a weight to those values depending on how far in the future that moment lies, and then sum up those weighted values over all moments in the future. This gives my utility for the whole outcome, and that's what's fed into my decision-making—it gives the utilities whose expectation standard decision theory tells me to maximize. If I apply smaller weights to later moments, I am discounting the future.

Is such discounting always unreasonable? This is a question over which there is heated debate among philosophers, psychologists, and economists. In the *Protagoras* (356a-e), Plato says it is irrational. He thinks it involves a cognitive error similar to the one we commit if we judge something further away to be smaller because it appears so. In his *Reasons and Persons*, on the other hand, Derek Parfit says it is permissible (Parfit, 1984, 313). He thinks we identify less with the person we'll be in the far future and so care less about them acquiring goods. And even those who agree that some amount of discounting is permissible often disagree about how much it is reasonable to discount (Frederick et al., 2002). Perhaps it is reasonable to prefer one bar of chocolate now to two bars in a year, but not one bar now to five thousand bars tomorrow; having the latter preference would require discounting the future too much.

Now suppose that the nudger and I disagree over whether I am unreasonable to discount the future as much as I do when I prefer not to contribute to my pension scheme. What then? Is it legitimate for the nudger to present the options in a way that mobilises certain psychological mechanisms that push me towards contributing? Economists have presented arguments that certain patterns of discounting at different times is irrational, but among those patterns that are not affected by these arguments, there are ones that discount the future very heavily indeed. And so the nudger cannot appeal to those arguments. They must instead ground their prescriptive theory of means-ends rationality in their own intuitive judgments of what is rational or reasonable, and what is not. And we have no neutral

ground from which to adjudicate the dispute between me and the nudger. And so the nudger's prescription looks paternalist in a way that the liberal should reject.

And this concern arises not only for different views about reasonable rates at which to discount the future. The same goes for people who disagree about what attitudes to risk are reasonable. Some will think that only risk-neutral attitudes are permitted; others will permit more, but they will identify the limits of reasonable risk-aversion and risk-inclination differently. What if nudger and nudged disagree over whether the risk attitudes of the nudged are reasonable or not?

And so our third question to the liberal is this: If there is disagreement among citizens about the correct theory of rational choice and the bounds of reasonableness within that theory, on which theory should a government base its policy-making? And, in particular, when is it legitimate for the government to nudge citizens to choose against their own preferences on the grounds that those preferences are the product of an unreasonable theory of rational choice?

5 Conclusion

Three complaints we hear often about liberalism at the moment: (i) it lacks, or actively prevents, the means by which to transform the societies we actually live in now into the society the liberal would most like to see; (ii) favouring a sort of neutrality between a great pluralism of competing views, it lacks the means by which to adjudicate in cases of profound moral disagreement; and (iii) it is prone to treat those whose views lie too far from the views of those in positions of power as unreasonable and therefore fails to represent their true interests. Are these part of the catalyst for the new illiberalism? I don't know. But they are important challenges that liberalism needs to meet. I have tried to spell out some of the philosophical issues surrounding these challenges.

References

- Bernoulli, Daniel (1738 [1954]). Exposition of a New Theory on the Measurement of Risk. *Econometrica*, 22(1), 23–36.
- Bottomley, Christopher, & Williamson, Timothy Luke (ta). Rational risk-aversion: Good things come to those who weight. *Philosophy and Phenomenological Research*.
- Buchak, Lara (2013). *Risk and Rationality*. Oxford, UK: Oxford University Press.

- Carney, D. R., & Banaji, M. R. (2012). First is Best. *PLoS ONE*, 7(6), e35088.
- Chew, Soo Hong (1983). A generalization of the quasilinear mean with applications to the measurement of income inequality and decision theory resolving the Allais paradox. *Econometrica*, 51, 1065–1092.
- Chew, Soo Hong (1989). Axiomatic utility theories with the betweenness property. *Annals of Operations Research*, 19, 273–298.
- Delmas, Candice (2018). *A Duty to Resist: When Disobedience Should Be Uncivil*. Oxford, UK: Oxford University Press.
- Frederick, Shane, Loewenstein, George, & O'Donoghue, Ted (2002). Time Discounting and Time Preference: A Critical Review. *Journal of Economic Literature*, 40(2), 351–401.
- Gaus, Gerald (2010). *The Order of Public Reason*. Cambridge, UK: Cambridge University Press.
- Parfit, Derek (1984). *Reasons and Persons*. Oxford University Press.
- Quiggin, John (1982). A Theory of Anticipated Utility. *Journal of Economic Behavior and Organization*, 3, 323–43.
- Quiggin, John (1993). *Generalized Expected Utility Theory: The Rank-Dependent Model*. Kluwer Academic Publishers.
- Quong, Jonathan (2023). Public Reason. In E. N. Zalta (Ed.) *Stanford Encyclopedia of Philosophy*. Metaphysics Research Lab, Stanford University.
- Rawls, John (1999 [1975]). *A Theory of Justice (Revised Edition)*. Cambridge, Mass.: Harvard University Press.
- Sunstein, Cass R. (2023a). On Liberalism. Unpublished manuscript. Available at: https://papers.ssrn.com/sol3/papers.cfm?abstract_id=4644402.
- Sunstein, Cass R. (2023b). Why I Am A Liberal. *The New York Times*. 20th November 2023. Available at: <https://www.nytimes.com/2023/11/20/opinion/cass-sunstein-why-liberal.html>.
- Thaler, Richard H., & Sunstein, Cass R. (2003). Libertarian Paternalism. *American Economic Review*, 93(2), 175–79.
- Thaler, Richard H., & Sunstein, Cass R. (2008). *Nudge: Improving Decisions about Health, Wealth, and Happiness*. New Haven: Yale University Press.
- Thoma, Johanna (2023). Merely Means Paternalist? Prospect Theory and 'Debiased' Welfare Analysis. *Philosophy of Science*.