

Autonomy's Substance

FABIAN FREYENHAGEN

ABSTRACT *In this article, I argue that autonomy has to be conceived substantively in order to serve as the qualifying condition for receiving the full set of individual liberal rights. I show that the common distinction between content-neutral and substantive accounts of autonomy is riddled with confusion and ambiguities, and provide a clear alternative taxonomy. At least insofar as we are concerned with liberal settings, the real question is whether or not the value(s) and norm(s) implied by an account of autonomy are acceptable to reasonable people, not whether these accounts are content-neutral, procedural or input-focused. Finally, I demonstrate how substantive constraints are compatible with, or even implied in, the notion of autonomy at play in (Rawls's) political liberalism. Overall, I present a normative reconstruction, clarification, and internal critique of liberalism, drawing on case law and statutes from England and Wales.*

1.

It has become a commonplace to distinguish between content-neutral and substantive accounts of autonomy.¹ Often, this distinction is treated as synonymous, or at least partially overlapping, with the distinction between procedural and substantive accounts.² Usually these respective distinctions are deployed in the context of arguments to the effect that substantive accounts are not fit for the purpose of fixing the threshold at which adults are granted the full set of liberal rights. Such accounts may serve other functions (providing an ideal of character or a rallying cry for social emancipation and feminist struggle), but they are — so the argument goes — unsuitable for this key function which accounts of autonomy are meant to perform in liberal settings.³

In this article, I argue that autonomy *has to be conceived substantively* if it is to serve as the condition which fixes the threshold at which someone is granted the status of having the full set of liberal rights.⁴ (It is autonomy in that sense alone that I will be concerned with — not autonomy as a property of an ideal character or a political entity, but as a threshold concept for fixing a legal status, where this requires possessing certain capacities.) I also argue that the common distinction between content-neutral (and procedural) accounts, on the one hand, and substantive accounts, on the other hand, is riddled by confusion and ambiguities. I suggest that, at least insofar as we are concerned with liberal settings, the real question is instead whether or not the value(s) and norm(s) implied by an account of autonomy are acceptable to reasonable people, not whether they are content-neutral or -partisan, procedural or non-procedural, input- or output-focused.

My aim is a normative reconstruction, clarification, and internal critique of liberal practices and theories, not a defence of them against non-liberals (or even an endorsement of them). I am particularly concerned with those contemporary liberal theories that

set out the normative potential of modern liberal democratic societies, such as notably the later Rawls and those working within his framework. Looking at actual practices and institutions is informative for such theorising (albeit not determinative of its content), since it reveals the historical background of such theorising and what it takes to implement its results. My examples will be from England and Wales, but other liberal jurisdictions face similar issues and challenges — the chosen example context merely brings these better to the fore.

The article is structured as follows. First, I lay out the motivation for content-neutral accounts in more detail from the perspective of those who advance these accounts — at this point, I am not yet evaluating whether these motivations suffice to justify the focus (Section 2). Then, I start chipping away at them and at the various ways in which the distinction between content-neutral and substantive accounts of autonomy is drawn. I distinguish three ways in which this distinction can be drawn and show that the strongest claim to content-neutrality — namely eschewing all normative content at all levels (the outcome of the decision, reasons for the decision, and competencies required for its being autonomous) — is unworkable (Section 3). Next, I also demonstrate that the two other versions of content-neutrality — neutrality about the contents of the decisions and the reasons for it — are not well suited to their supposed purpose (Section 4). Finally, I demonstrate how substantive constraints are compatible with, or even implied in, the notion of autonomy at play in (Rawls's) political liberalism (Section 5).

2.

What purpose does talk of 'content-neutral' (including 'procedural') accounts of autonomy, in contrast to 'substantive' accounts, serve? The initial answer one is likely going to get is that it is meant to serve the purpose of providing an account of autonomy that does *not* give rise to unwarranted paternalism by the state (or state-authorised agents).⁵ Paternalism is sometimes warranted — many ways in which parents (are supposed to) treat their children would be examples. However, it is not warranted in regard to an autonomous adult's making a self-regarding decision, even if it is an unwise decision. Critics of substantive accounts of autonomy claim that such accounts will lead to unwarranted paternalism because they are perfectionist,⁶ whereby perfectionism is understood as the doctrine that there is one true and objective ideal of the good life that the state ought to promote and implement. Perfectionist accounts of autonomy build the conception of the truly good life into the account of autonomy itself, such that one is not (fully) autonomous — on these accounts — if one does not live the good life (and one is not even capable of autonomy if one is not capable of living such a life).⁷ The problem with perfectionism is that it is incompatible with reasonable pluralism — that is, it is incompatible with the fact that people willing to be convinced by arguments continue to disagree about fundamental questions in ethics (and philosophy more generally) even after free and extended discussion.⁸ If perfectionism is nonetheless used as government policy, then it becomes tyrannical in its attempts to 'force people to be free'.

In sum, for liberal political theorists (particularly those working within the framework of the later Rawls), the ultimate point of contrasting 'content-neutral' (including 'procedural') accounts of autonomy with 'substantive' ones is to ensure that *those accounts are distinguished and then excluded that are incompatible with reasonable pluralism*.⁹ Only in this

way — the argument goes — can one avoid the unwarranted paternalism that perfectionism would licence. A content-neutral account can 'accommodate a diversity of desires and ways of life as autonomous' and thereby 'fits well with standard accounts of political liberalism'.¹⁰

The underlying concern relates to a long history, reaching back at least as far as the early modern wars of religion and the subsequent attempts to secure freedom of conscience and religion. One high point in this history is Mill's famous claim that:

The only freedom which deserves the name, is that of pursuing our own good in our own way, so long as we do not attempt to deprive others of theirs, or impede their efforts to obtain it. [. . .] Mankind are greater gainers by suffering each other to live as seems good to themselves, than by compelling each to live as seems good to the rest.¹¹

Mill's claim has echoes in contemporary legal cases. Consider, for example, this extract from a judgement by LJ Donaldson:

An adult patient who, like Miss T., suffers from no mental incapacity has an absolute right to choose whether to consent to medical treatment, to refuse it or to choose one rather than another of the treatments being offered. [. . .] This right of choice is not limited to decisions which others might regard as sensible. It exists notwithstanding that the reasons for making the choice are rational, irrational, unknown or even non-existent.¹²

This piece of legal rhetoric brings out clearly not just the idea of neutrality between conceptions of the good, but even what might be called '*reason neutrality*' — there is no restriction whatsoever on the reasons which a competent person may have for deciding as s/he does (they can be 'rational, irrational, unknown or even non-existent').¹³ There is neutrality even regarding norms of justification, not just regarding the conceptions of the good invoked in justifications.

Ultimately, the idea is to protect individuality and eccentricity, as the following passage from a recent case (concerning the making of a statutory will) brings out nicely:

Just as a testator has always had the freedom . . . to make testamentary dispositions which are unreasonable, foolish or contrary to generally accepted standards of morality, so too a person in his lifetime has the freedom to act in a manner which is (for example) unwise, capricious, or designed to spite his relations.¹⁴

When it comes to their own affairs (including, as in this case, the making of one's last will),¹⁵ autonomous persons should be able to make decisions which others consider unwise, or even 'capricious'.¹⁶

One question that arises here is whether any autonomy-based liberalism is going to be compatible with reasonable pluralism, given that autonomy is, arguably, itself a contestable value. If autonomy is, indeed, itself a value about which reasonable persons disagree, then it is not suitable to underpin, or be part of, a liberal system which is meant to be justifiable to all of its (reasonable) citizens. For example, consider again the final sentence quoted from Mill above: 'Mankind are greater gainers by suffering each other to live as seems good to themselves, than by compelling each to live as seems good to the rest'. Mill is a champion of individual liberty, but he justifies this by claiming that it is the

best means to securing the long-term wellbeing of humanity, which in turn he understands in terms of utility. This justification is contestable — one could argue (perhaps with a view to the environmental degradation and destruction of the last 150 years) that individual liberty is not the best means to the long-term-wellbeing of humanity; one could object to the Utilitarian notion of wellbeing deployed by Mill; or one could counter that, say, the dignity of some should not be sacrificed to the long-term wellbeing of humanity (such that one should, for example, outlaw ‘dwarf-tossing’ or ‘laser-quest’ or trade in human organs, even if people voluntarily engage in them and humanity’s long-term wellbeing would be served by allowing them).

It is remarkable how little theorists who adopt a content-neutral account of autonomy discuss this consideration. They tend to go on as if this wider issue is unimportant and the only thing that matters is the more downstream question of how autonomy is internally configured, such that it is compatible with different conceptions of the good. This rests on two assumptions: (1) that the so configured autonomy can be defended without, more upstream, invoking a contestable conception of the good; and (2) that the only thing one can (reasonably) disagree about is the (conception of the) good. Neither assumption is warranted — the above little foray into Mill’s (upstream) justification of autonomy suggests that (1) is far from guaranteed; and the fierce way in which especially the debate about abortion (which, I take it, concerns rights and justice, not primarily the good) and also meta-ethical debates are conducted, suggests that reasonable disagreement persists in such matters too.¹⁷

Some do consider the issue of whether autonomy-based liberalism is compatible with reasonable pluralism. They then either bite the bullet and accept that autonomy is a contestable value among reasonable persons, albeit one that internally requires compatibility with different conceptions of the good.¹⁸ Or they argue that liberals should adopt a ‘political’ conception of autonomy — one that can be part of an overlapping consensus of reasonable worldviews, and that is not metaphysical or comprehensively ethical.¹⁹ They suggest that only content-neutral accounts could fulfil this role. We will see later that this is mistaken, but for now I just want to note that something like a move to a political conception of autonomy is required if one’s ultimate purpose is compatibility with reasonable pluralism.

3.

Having reviewed *the point* of contrasting ‘content-neutral’ (and ‘procedural’) accounts of autonomy with ‘substantive’ ones, I now discuss *how the contrast is drawn*, bringing out some ambiguities and confusions. Also, I show in this section that the suggestion that there could be accounts of autonomy that are devoid of normative content is mistaken.

Let us begin in illuminating the contrast between ‘content-neutral’ and ‘substantive’ accounts of autonomy by looking at those accounts that are presented as falling squarely in the former camp. Gerald Dworkin brings out the ‘content-neutrality’ of his view as follows:

In my conception, the autonomous person can be a tyrant or a slave, a saint or sinner, a rugged individualist or champion of fraternity, a leader or follower.²⁰

Christman echoes this when drawing out an implication of his account:

One implication of this theory is that people could turn out to be autonomous despite having desires for subservient, demeaning, or even evil things and lifestyles. I don't take this to be a defect of the model. It only reveals that the conception of autonomy we are discussing is 'content neutral'.²¹

This suggests that the contrast is to those accounts of autonomy which are, for want of a better term, *content-partisan*, and 'substantive' should thus be understood in this way. Surprisingly, a good example of a content-partisan account of autonomy is actually Mill's account, according to which we cannot autonomously (or, as he has it, 'freely') choose to sell ourselves into slavery or enter (other) irrevocable contracts.²² Such restrictions on what kind of decision outcomes one can autonomously choose are actually quite rare in the literature, but real life might provide other examples, such as social workers who judge that one cannot autonomously decide to live in squalor.²³

So far, it would seem that the contrast between content-neutral and content-partisan accounts of autonomy is one between *outcome-neutral* and *outcome-partisan* accounts. The substantive views are characterised such that certain decisions — such as being a slave, being capricious to one's relatives, or living in squalor — cannot be the outcome of autonomous decision-making, whereas content-neutral accounts permit any decision to be the outcome of autonomous decision-making. In legal contexts, this distinction is sometimes marked as one between 'functional' and 'outcome' tests of decision-making capacity (or competence).²⁴

However, if we recall LJ Donaldson's judgement in *Re T* [1992] and consider various statements in the philosophical literature, we see that content-neutral accounts are actually often more demanding than merely requiring outcome-neutrality. Such statements suggest that we should also reject any views which restrict the content of the reasons one has for choosing a certain option. Thus, a *reason-partisan* account of autonomy might deny that one can autonomously decide to commit suicide for the reason of minimising physical pain, but accept that one can autonomously decide to commit suicide for some other reason.²⁵ Such an account would be outcome-neutral in that no specific decision outcome would be ruled out in principle as non-autonomous, but might still be described as content-partisan, since it is not neutral about the content of the reasons (or, if that is a wider category, motivations) that can be deployed by an autonomous agent. And there could be a combination of the two content-partisan conditions I mentioned so far — some accounts might be outcome- and reason-partisan. Characterisations of content-neutral accounts of autonomy are not consistent in whether they exclude being outcome-partisan,²⁶ reason-partisan,²⁷ or both.²⁸ Let us speak of '*strong content-neutrality*' if both outcome- and reason-neutrality is required, i.e. when autonomy is defined as making '[. . .] no reference to constraints on the content of a person's choices or the reasons he or she has for them'.²⁹

The ambiguity of the term 'content-neutrality' does not stop here, however. Some authors use it in an even stronger sense still. To see this, consider how Benson distinguishes between different kinds of substantive views. Benson, who (at least since 1994) holds an account which is strongly content-neutral in the sense just introduced, describes how his account does *not* fit into the standard carving up of the terrain:

I sketch part of an alternative conception of autonomy that incorporates normative content but does not constrain *directly* the types of actions agents might autonomously perform or the content of the motives or values that lead them to act.³⁰

This suggests an important distinction: some normatively laden accounts of autonomy *constrain directly* the decision outcomes and/or reasons for them, while other such accounts constrain the decision outcomes and/or reasons *merely indirectly*. Directly substantive views are outcome-partisan and/or reason-partisan, but indirectly substantive views are neither; and, if process is understood in a sufficiently wide sense, the latter are even procedural.³¹ The following view is an example:³² an agent has to take herself to have the authority to decide for herself, but this is, at least in principle, compatible with choosing all sorts of outcomes (including suicide, slavery, or living in squalor) and all sorts of reasons for choosing them (including the desire to be a servile wife to one's husband). This also means that indirectly substantive accounts allow for the possibility of unwise yet autonomous decisions:

The normative competencies these theories describe, however, need not entail any direct, normative restrictions on the contents of autonomous agents' preferences or values. These theories typically allow that normatively competent persons can choose what is unreasonable or wrong or value what is bad, because competence lies some distance short of perfect evaluative perception or responsiveness.³³

Still, indirectly substantive accounts of autonomy are normatively substantive insofar as the competencies in question cannot be characterised correctly in a purely descriptive way.³⁴ Indeed, Benson speaks of these accounts in contrast to 'content-neutral' (and directly substantive) ones.³⁵ This only makes sense if he uses 'content-neutrality' in an *extra strong* sense: as excluding all normative content, whether at the level of decision outcome, reasons for the decision, or competencies involved in making it. Thus, in his hands, the distinction between content-neutral and substantive views turns into a distinction between accounts of autonomy *without any normative content* and those with some such content.

If not already before, then at least at this point, the reader should feel like pushing back against the purported distinction. None of the supposedly content-neutral accounts of autonomy could avoid invoking some normative content in characterising the competencies or procedures it requires. Thus, Dworkin lists the ability for critical reflection and procedural independence as his supposedly content-neutral conditions on autonomy. Neither can be captured in norm-free ways. For example, procedural independence excludes influences such as manipulation or coercive persuasion, and these, I take it, are ineliminably normative notions. Critical reflection is an ability, which, *qua* ability, has an in-built standard that one can fall short of, or even fail to meet entirely; and to make the possession of this ability a condition of autonomy, is to endorse the (contestable) judgement that this norm should be authoritative in a certain context. Similarly, Christman lists an authenticity (or to be precise non-alienation) condition along with a basic decision-making capacity (encompassing the absence of psychopathological influences on the reasoning process as well as of manifest inconsistencies). Again, we could not even make sense of these as *abilities* of, and thereby *constraints* on autonomous

agency, if we conceived of them in a completely norm-free way. Indeed, Christman admits as much when he writes:

Defining a 'constraint' necessarily involves specifying (if only implicitly) the range of human actions and pursuits that such constraints make impossible [...] The very meaning of constraint presupposes a range of normal (and perhaps morally valued) action types that humans are thought to pursue.³⁶

Something similar holds true of procedures too. It might be often overlooked, but procedures contain (at least implicitly) normative substance.

Consider, for example, the choice between a first-past-the-post and a proportional representation electoral system: each of these procedures connects to, and in a certain sense embodies, different norms, such as clear accountability and a decisive outcome in the case of first-past-the-post and a closer relationship between the votes cast and the parliamentary make-up in the case of proportional representation. Similarly, there are a range of procedures to decide criminal or civil actions — for example by way of a due-process trial or by way of combat or ordeal — and which one we choose depends, at least in part, on the normative content contained in these processes.

Returning to decision-making procedures, it is striking how normatively laden the descriptions are that are employed by self-described proceduralists, such that there is talk of an agent's having an '[...] *accurate* conception of herself [...]' and '[...] *normal* cognitive functioning';³⁷ of her '*deliberating correctly*';³⁸ of being '*appropriately* sensitive to evidence'.³⁹ Reasoning, as Buss writes, is a '*norm-governed process*'.⁴⁰ Requiring of persons that they understand and weigh in the balance all the relevant information, including future consequences, for every decision would be to impose a particular way of life on them — the examined way of life, which could be accused of being overly intellectualist.⁴¹ Even demanding that a decision-making process is such that it does not involve any manifest inconsistencies,⁴² is to import normative content: consistency is just one norm among others, and one that is contestable, as, for example, the possibilities of true moral dilemmas and of moral conversions suggest. If I really ought to care for my ailing mother *and* really ought to go and join the resistance, and I cannot do both, then my decision-making process is not deficient for involving this manifest inconsistency. Equally, Paul's decisions might be manifestly inconsistent with Saul's worldview, but this does not necessarily mean that they are faulty or any less autonomous.

This shows that there can be no extra strong content-neutral accounts — normative content will have to come in at some level, such as, notably, the level of the competencies or procedures required for a decision to be autonomous. This leaves only a distinction between directly and indirectly substantive views. Before I suggest that even this distinction is not really where the action is, I consider one objection to the argument just made.

One might argue that, while it is, indeed, impossible to eschew all normative content, it is at least possible to eschew specifically ethical content. Thus, content-neutral accounts of autonomy would be unavoidably value-laden in the competencies and processes on which they insist (and have to insist), but the values in question would be epistemic (concerning, for example, accuracy and sensitivity to evidence), not ethical. As such, a real contrast to perfectionism — which by definition includes ethical content: requirements as to what makes human life go well or even best — would remain and

thereby it would continue to make sense to distinguish between content-neutral and substantive (and as such perfectionist-prone) accounts of autonomy.

To argue in this way is to overlook two related points that are important in our context. First, there is reasonable pluralism *also* about epistemic issues, such as regarding the question of what counts as sensitivity to evidence (and, indeed, what counts as evidence). Hence, as soon as normative content — even supposedly non-ethical, mere epistemic one — is at issue, the question of compatibility with reasonable pluralism is at issue, and notions of autonomy are not simply made compatible by being (supposedly) non-ethical. Moreover — and this is the second reason — it is not even clear that one could draw a neat distinction between what is ethical and non-ethical in our context. Reasonable pluralism extends as to where this distinction ought to be drawn. Thus, for example, a thorough-going perfectionist would want the state to make us also epistemically (and aesthetically) better, not just in terms (or because) of our generosity towards others or courage in the face of adversity. For such a perfectionist, the ethical would encompass all of the normative, directly or at least indirectly. While the perfectionism would presumably be incompatible with reasonable pluralism, such a view of the scope of the normative seems clearly compatible — it seems that some persons willing to be convinced otherwise will maintain the broad view of the ethical even after free and extended discussion. At any rate, the possibility of drawing the distinction between the ethical and non-ethical normative in different reasonable ways (or even denying that distinction altogether) makes it unsuitable for the purpose of delineating those accounts that are compatible with reasonable pluralism from those that are not.

4.

Earlier I spoke of ‘legal rhetoric’ in referring to LJ Donaldson’s judgement in the *Re T* case. I did this for two reasons. Firstly, it is a striking fact — and no small irony — about this case and similar cases that judges tend to affirm state neutrality and the absolute rights of persons to decide in their personal affairs particularly when they have actually found someone to lack the requisite decision-making capacity to qualify for the full set of liberal rights. Someone ‘like Miss T’, but crucially not Miss T, has an absolute right of refusal attested by LJ Donaldson in the *Re T* case. Indeed, some commentators note that judges smuggle in value judgements — in my terms, outcome- and reason-partisanship — into such decisions.⁴³ Secondly, and perhaps more importantly for our context, I spoke of ‘legal rhetoric’ since, strictly speaking, it is false what LJ Donaldson said. Even competent persons have no absolute right regarding treatment decisions, for they may be hospitalised and treated under mental health legislation (in England and Wales, to stick with our example context, under the *Mental Health Act 1983*, as amended 2007).⁴⁴ Such legislation is risk-based — specifically risk of harm to oneself (or others). Someone may be competent regarding a decision, but he or she would still be involuntarily treated, if the risks of self-harm are sufficiently large and these risks arise in the context of a mental disorder of a particular nature and extent.⁴⁵ Insofar as competence is the legal marker of autonomy, this means someone’s autonomy can be overridden in practice to protect this person from self-harm. Here, competent adults are interfered with not just to protect the freedom (and rights) of others (as Mill’s well-known ‘harm principle’ allows), but also to protect them for their own good (as this principle rules out in the case of competent adults).

What both of these points bring out is that liberal legal settings, *as a matter of fact*, are not content-neutral in either an outcome- or reason-neutral sense (never mind free of all normative content). Now, this by itself does not yet settle the matter as to autonomy's substance. It might be that these considerations about the actual practice of legal regimes merely tell us that such regimes have still a way to go before living up to their liberal aspiration. Both the letter of the law and the spirit of its implementation would have to be reformed — and as a consequence people like Miss T will have their autonomy recognised and mental health legislation will become competence-based, not risk-based. If we want to know about liberal aspirations, then we do better to turn to liberal theories than liberal practices. These theories have delineated the proper ideal liberals should aspire to achieve. In this way, one might reject the import and relevance of any appeal to actual practices.

However, this is not the best reaction for liberals. For a start, liberal practices are actually as (or even more) revealing about liberal considered judgements — the building blocks for theorising about ideals — than what theorists propose counterfactually in the rather shielded environment of their studies and seminar rooms. For all the failings and biases of the legal system, what withstands the scrutiny of courtrooms should at least deserve a hearing too. Its actors (at least in the common law tradition on which I am basing my account here) employ an adversarial method that is well-suited to bring out strengths and weaknesses on both sides (especially when the decisions are the subject of appeal and make their way through a number of court levels); and these actors draw on a long series of precedents concerning a wide range of complex, real world cases. Indeed, given their higher-order commitment for reflectiveness of the intuitions that are the building blocks for theorising, liberal theorists should give more weight to actual liberal practices than they tend to realise. What matters for them are not raw intuitions, but considered judgements refined by way of seeking wide reflective equilibrium. Surely, the insights from actual legal practice are part of this process. True, the insights reflect the actual existing practices and there is a danger of a *status quo* bias; but the way we engage in counterfactual reasoning and imagination, at least indirectly, also reflects existing practices and also involves the danger of *status quo* bias — indeed, it might be a safer response to this bias to tackle one's entanglement with the existing social world directly, rather than have recourse to (what one takes to be) ideals and fall into the trap of thinking that subject matter alone (thinking about ideals rather than existing practices) or engaging in counterfactual reasoning suffice to escape this entanglement.

At least some liberal theorists express the view that they want to build on existing practices, rather than (pretend that they can) theorise in a void. They acknowledge that the content, scope and justification of their theories will depend on the nature of the practices to which they apply; in short, they advance 'practice-dependent' theories.⁴⁶ While the existing practices and institutions (and their history) are not (meant to be) solely determinative for how liberal theories are practice-dependent, they are the starting point and are ignored at liberalism's peril. At any rate, assigning weight to actual practices promises to clarify liberal practices and theories, albeit — as I am the first to admit — not offer a ground-up defence of them.

The more promising reply to the actual cases of restrictions on individual's choices and reasons for them would be to argue that these cases merely bring out that there are other values — such as the preservation of human life — we have to combine with a

content-neutral idea of autonomy in order to have the most defensible liberal state, and this might mean that sometimes (content-neutral) autonomy has to be overridden. This might be what is at issue in risk-based mental health legislation and the way the question of competency (and thus autonomy) is either completely disregarded or not central to this legislation, in effect trumping autonomy (in those cases where it is present) with the protection of (mental) health and life. Often enough, real legal cases and policies underdetermine whether their justification is based on autonomy-internal considerations (involving, say, an outcome-partisan conception of autonomy) or based on other consideration or both.

However, it is possible to argue for the same conclusion without relying directly on appeals to the actual reality of liberal legal regimes. Instead, a more conceptual point about their nature can do the argumentative work. Compatibility with reasonable pluralism can be achieved by reason-partisan accounts autonomy, which constrain the contents of the reasons which an autonomous person can invoke in choosing, or even outcome-partisan accounts, which directly constrain the content of what can count as autonomous decision. It all depends on whether these constraints on what can be autonomously chosen and the reasons for choosing it, would be acceptable to reasonable persons — and whether they are so acceptable is not something that could be pre-judged, as it were, *a priori*, but would depend (in Rawls's terms) on the 'public political culture' of the society in question. It might be that in a given society, an overlapping consensus emerges such that, for example, certain contracts — perhaps involving very high interests rates or the risk of debt bondage — are seen as something no one could autonomously agree to, and, hence, in such a society the state would not enforce such contracts. Or it might be that in such a society an overlapping consensus exists such that parents cannot exclude their children from their last will simply because of dislike or callousness — not because the detrimental effects this might have on values other than autonomy, but as not properly expressing autonomy (which, in this society, is understood to contain internally some moral constraints, perhaps because the self that is doing the legislating and determining is understood in a minimally moral way, such that people could not recognise themselves in outright callousness). Such constraints on autonomy might lead to much less interference (as hardly anyone actually wanted to make callous wills or tried to enter the contracts in question). It might also be compatible with more reasonable worldviews than insisting, for example, that people actually follow certain prescribed deliberative procedures (something certain communities, such as religious ones which insist on the passivity of human decisions as part of an openness to their god or gods, might reasonably reject as too demanding). Hence, unless convincing arguments are supplied why such constraints would be in principle incompatible with reasonable pluralism, insisting on outcome- and reason-neutral accounts of autonomy is not serving the liberal purpose. Rather, liberals should ask a different question and then distinguish accounts according to the answer to that question: *is the proposed account of autonomy, whether it be directly or indirectly normative, compatible with reasonable pluralism?*

What matters, then, is *not* whether or not the account is content-neutral in any of the three senses delineated, makes merely procedural or structural requirements, and focuses only on the inputs into the decision-making process; but whether or not citizens of a body politic who are willing to be convinced otherwise would agree on it after free and extended discussion.

5.

One might object that the mere possibility that an overlapping consensus according to which autonomy includes reason-partisan or even outcome-partisan constraints could emerge is insufficient to prove my point. What would be needed is evidence of actual agreement of the sort that would suggest such an overlapping consensus about reason-partisan and/or outcome-partisan features of autonomy. Indeed, even this will not suffice — for actual agreement about what is compatible with reasonable pluralism might not anticipate an outlook that an actual or potential member of the community might hold in the future, where such outlook is denying the features in question in a way that is reasonable (i.e. this person denies them after free and extended discussion, even though s/he is willing to be convinced otherwise).

In response, the first point is that it is a genuinely difficult challenge to demonstrate that a (consensus) position is compatible with reasonable pluralism. However imaginative we might be, history tells us that we human beings are not particularly skilled at anticipating normative outlooks that call into question the reigning consensus but are nonetheless reasonable. Various positions regarding slavery or women might well have been consensus position at the time and were then seen as irrevocable, but we think of them now as unreasonable — as something to which no one could agree to after free and extended discussion, even if willing to be convinced otherwise.

Yet, the fact that judgements about what is (and is not) compatible with reasonable pluralism are preliminary and fallible is not a specific problem for what I propose here, but inherent to the (Rawlsian) framework I internally criticise and expand. One need only recall Rawls's own approach to demonstrate the compatibility of his political liberalism with reasonable pluralism: he engages in a series of case studies of worldviews that have survived the test of time and are in common currency today in order to show that each of them is compatible with political liberalism.⁴⁷ Even if he were successful in these case studies — something that is rather questionable⁴⁸ — it should be clear that this could only give rise to a preliminary and fallible result.

In line with my strategy of internal critique, then, it would suffice to take Rawls at his word that political liberalism is compatible with reasonable pluralism and then to demonstrate how a reason-partisan and/or outcome-partisan conception of autonomy is implied by, or at least compatible with, political liberalism. The latter compatibility would imply the former and, in that sense, also imply non-perfectionism (as conceptualised within this framework).

I cannot hope to offer a full demonstration here, so two examples will have to suffice. First, courts, regulative bodies, and parliaments have, particularly in recent years, restricted loan and mortgage arrangements in ways that suggests a restriction of decision outcomes on the basis of (a certain conception of) autonomy. In particular, they have capped the maximum interest to be charged for such financial instruments and/or made it a requirement on financial institutions that they only sign up those customers who pass affordability tests.⁴⁹ These measures have broad support across party lines and class divisions, and from among very different worldviews. The effect of these regulations is that customers cannot — in a legally binding fashion — choose to enter contracts that are deemed to charge excessive interest rates or are unaffordable for them even if the interest rates are not excessive. This restriction can be made sense of in Rawlsian terms. The idea of citizens as free and equal is a central part of the overlapping consensus for

Rawls. It includes a conceptualisation of autonomy in terms of two moral powers — the capacity to form, act on, and revise a conception of the good and the ability to be guided by a sense of justice.⁵⁰ Protecting autonomy in this sense requires providing the fair value of the equal political liberties.⁵¹ Also, health-care provision is justified as a way of securing this autonomy: this provision has to be secured because sufficient health is a condition of possibility for exercising one's function as equal citizen and for keeping one's status as equal citizen. In a structurally similar way, limits on who can get debts and caps on interest rates can be justified: getting caught in an unsustainable level of debt or being vulnerable in a way that one would accept interest rates that require one to pay back more than double of what one borrowed is so detrimental to one's functioning and standing as equal citizen that the state has to take steps to prevent them from occurring — otherwise one's two moral powers and thereby autonomy would be too curtailed. Even more clear-cut cases than these current policy issues would be slavery and debt bondage — contracts involving these would be null and void on the same Rawlsian grounds, adding further constraints on what would count as autonomous decision outcomes. As a result, Rawlsian autonomy is outcome-partisan at least in respect with certain decision outcomes (such as the decision to take out unaffordable loans or to sell oneself into slavery or to offer vital organs as collateral for loans). And yet — as granted for the sake of internal critique — it is compatible with reasonable pluralism and as such perfectionist. (Indeed, there is a further reason for thinking that such outcome-partisan autonomy would not be perfectionist: this conception of autonomy would be a requirement of justice, not about promoting the good, and perfectionism is normally understood to consist in promoting the good, not securing justice.) Put in a slogan, the later Rawls' conception of autonomy is *both* 'political but not metaphysical' *and* substantive.

Second, we can confirm this by thinking about inalienable rights. There might be a variety of reasons for such rights, but, at least in some instances, the justification seems to be autonomy-based and to betray an outcome-partisan conception of it. For example, the closest thing to a constitution in Germany — the *Grundgesetz* [Basic Law] — protects freedom of association in a way that makes this protection inalienable by individuals, particularly when it comes to forming unions and professional organisations:

The right to form associations to safeguard and improve working and economic conditions is guaranteed to everyone and to all trades and professions. Agreements which restrict or seek to hinder this right are null and void; measures directed to this end are illegal.⁵²

This makes, among other things, labour contracts that include anti-unionisation clauses legally non-binding ('null and void'). Insofar as freedom of association is understood to be part and parcel of liberal autonomy and insofar as normally entering into contract is counted among the liberal rights to which passing the threshold of autonomy provides us access, this would suggest that autonomy itself rules out certain decision-outcomes (the signing of such labour contracts) as non-autonomous. Indeed, such a view is compatible with, even seems to be implied by, (Rawlsian) political liberalism. As briefly indicated above, the notion of autonomy at play in political liberalism is about our ability to function as free and equal citizen. This ability is endangered if our freedom of association is alienable, such that, for example, economic pressure could make us agree to restrictions with whom we associate. This seems to be a case where our co-authorising of the constitutional essentials binds us collectively in a way to ensure that we are never put in

the position of having to compromise our basic liberties. Specifically, the requirement of protecting the freedom to form unions might be important not just for securing better wages, but also for preventing subjection to too much control by others in our working environment, which, in turn, can undermine our ability to function as free and equal citizens.⁵³

6.

This article is an internal critique of liberal theory. I have demonstrated that the common distinction drawn between content-neutral and substantive conceptions of autonomy is ambivalent between weaker and stronger contrasts (Section 3) — the contrast is either to outcome-partisanship, or also to reason-partisanship, or to any normatively laden element at all (such as the competences on which autonomy is normally premised). I have also shown that, once we get clear about the three contrasts in question, the final contrast relies on the absurdity of thinking accounts of autonomy for fixing the threshold for liberal rights can be value-free and normatively neutral (Section 3). I argued that the other two do not track what liberals actually are concerned about — compatibility with reasonable pluralism (Sections 2 and 4). I then (in Section 5) substantiated my claim that reason-partisan and even outcome-partisan conceptions of autonomy are compatible with this ultimate concern of (political) liberalism by discussing a number of example restrictions and how they fit into the Rawlsian framework.

In sum, whatever merits liberalism might have, it has to own up to the normative substance of its conception of autonomy and adopt appropriate strategies in justifying its reign.⁵⁴

Fabian Freyenhagen, School of Philosophy and Art History, University of Essex, Wivenhoe Park, Colchester CO4 3SQ, UK. ffrey@essex.ac.uk

NOTES

- 1 See Section 3 below, with references. Substantive accounts tend to include content-neutral conditions on autonomy, so the issue is whether the latter suffice for autonomy or whether something substantive is necessary as well.
- 2 See, for example, Paul Benson, 'Feminist intuitions and the normative substance of autonomy' in J.S. Taylor (ed.) *Personal Autonomy* (Cambridge: Cambridge University Press, 2005), chapter 5, p. 126; John Christman, 'Relational autonomy, liberal individualism, and the social constitution of selves', *Philosophical Studies* 117 (2004): 143–164, at pp. 148, 151; 'Procedural autonomy and liberal legitimacy', in J.S. Taylor (ed.) op. cit., chapter 12, at p. 281; *The Politics of Persons* (Cambridge: Cambridge University Press, 2009), p. 172; John Christman & Joel Anderson, 'Introduction' in J. Christman & J. Anderson (eds) *Autonomy and the Challenges to Liberalism: New Essays* (Cambridge: Cambridge University Press, 2005), p. 3; Sigurdur Kristinsson, 'The limits of neutrality: Toward a weakly substantive account of autonomy', *Canadian Journal of Philosophy* 30,2 (2000): 257–286, at p. 266; and Catriona Mackenzie, 'Relational autonomy, normative authority and perfectionism', *Journal of Social Philosophy* 39,4 (2008): 512–533, at pp. 519f. There is no uniform usage of 'procedural' — most often it is seen as a sub-category, along with 'structural' views, of content-neutral accounts of autonomy (see, for example, Kristinsson op. cit., p. 266); but sometimes it is instead used as identical with the wider category of content-neutrality, which is then subdivided in other ways, such as into 'structural' and 'historical' accounts (Paul Benson, 'Free agency and self-worth', *The Journal of Philosophy* 91,12 (1994): 650–68, at pp. 653f). In this article, I focus on the overarching category of 'content-neutrality'.

- 3 See, for example, Christman 2009 op. cit., p. 175; on autonomy as a status marker for receiving the full set of liberal rights, see also Lubomira Radoilska, 'Introduction: Personal autonomy, decisional capacity, and mental disorder' in L. Radoilska (ed.) *Autonomy and Mental Disorder* (Oxford: Oxford University Press, 2012), pp. ix–xli, at pp. ix–x; Bert van der Brink, 'Political autonomy and agonistic citizenship', in: Christman & Anderson (eds), *Autonomy and the Challenges to Liberalism*, op. cit., chapter 11, at p. 251. On the different requirements of 'de jure political autonomy' and of other forms of autonomy, see also Mackenzie op. cit., p. 523.
- 4 Elsewhere, I put forward a further argument for this conclusion, pointing to the normative substance hidden in assuming that psychopathology is an obstacle to (content-neutral) autonomy (see F. Freyenhagen & T. O'Shea, 'Hidden substance: Mental disorder as a challenge to normatively neutral accounts of autonomy', *International Journal of Law in Context* 9,1 (2013): 53–70).
- 5 See, for example, Natalie F. Banner, 'Unreasonable reasons: Normative judgements in the assessment of mental capacity', *Journal of Evaluation in Clinical Practice* 18 (2012): 1038–44, at p. 1038; 'Can procedural and substantive elements of decision-making be reconciled in assessments of mental capacity?', *International Journal of Law in Context* 9,1 (2013): 71–86, here 71, 72, 84; Natalie F. Banner & George Szmukler, "'Radical interpretation" and the assessment of decision-making capacity', *Journal of Applied Philosophy* 30,4 (2013): 379–94, esp. p. 379.; Christman 2004 op. cit., esp. pp. 157f; 2009, op. cit., esp. pp. 135, 162, 185; see also Gerald Dworkin, *The Theory and Practice of Autonomy* (Cambridge: Cambridge University Press, 1988), esp. p. 10.
- 6 See, for example, Christman 2009 op. cit., p. 185.
- 7 Such a perfectionist account of *autonomy* should not be confused with Raz's perfectionist account of *the human good* — the latter is broader, and, while it constitutively includes autonomy, that notion is procedurally conceived.
- 8 John Rawls, *Political Liberalism* (New York: Columbia University Press, [1993], Paperback Edition 1996), Lecture II, §2.4.
- 9 One might object here that certain restrictions of content might be compatible with reasonable pluralism — indeed, I show this to be the case in subsequent sections. Here I am merely reporting how the argument for content-neutrality is supposed to run.
- 10 Christman & Anderson op. cit., p. 3.
- 11 John Stuart Mill, 'On liberty' [1859], published in J. Gray (ed.), *John Stuart Mill: On Liberty and Other Essays* (Oxford: Oxford University Press, 1991), pp. 1–128, at p. 17.
- 12 *Re T* [1992] EWCA Civ 18, §3; see also §§30, 37; available at <http://www.bailii.org/ew/cases/EWCA/Civ/1992/18.html> (last accessed 12 November 2013 at 16.18h).
- 13 See also *Sidaway v Board of Governors of the Bethlem Royal Hospital* [1985] UKHL 1, 904-905/p. 28; available at <http://www.bailii.org/uk/cases/UKHL/1985/1.html> (last accessed 12 November 2013 at 16.24h); and *Re MB* [1997] EWCA Civ 3093, §17(2); available at <http://www.bailii.org/ew/cases/EWCA/Civ/1997/3093.html> (last accessed 12 November 2013 at 7.49h).
- 14 *D v R (The deputy of S) & Anor* [2010] EWHC 2405 (COP); §39; available at <http://www.bailii.org/ew/cases/EWHC/COP/2010/2405.html> (last accessed 12 November 2013 at 16.38h).
- 15 Not all liberal states are as permissive as England when it comes to making one's last will — for example, Germany constrains this, such that children and other significant others cannot be wholly excluded from a will, but receive a 'Pflichtanteil [compulsory share]'.
 16 See also *Mental Capacity Act 2005*, §1.4 (available at http://www.legislation.gov.uk/ukpga/2005/9/pdfs/ukpga_20050009_en.pdf (last accessed 13 November 2013 at 9.15h)).
- 17 See F. Freyenhagen, 'Taking reasonable pluralism seriously: An internal critique of political liberalism', *Politics, Philosophy, and Economics* 10,3 (August 2011): 323–42. On disagreements about justice, see also Jeremy Waldron, *Law and Disagreement* (Oxford: Oxford University Press, 1999); and Timothy Fowler & Zofia Stemplowska, 'The asymmetry objection rides again: On the nature and significance of justificatory disagreement', *Journal of Applied Philosophy* 32,2 (2015): 133–46.
- 18 See, notably, Will Kymlicka, *Multicultural Citizenship: A Liberal Theory of Minority Rights* (Oxford: Clarendon Press, 1995), esp. chapter 8; and *Contemporary Political Philosophy: An Introduction* (Oxford: Oxford University Press), chapter 6.
- 19 See Rawls 1993/1996 op. cit., Lecture II, §6. This seems to be Christman's current approach (2009 op. cit., esp. pp. 6, 16, 114ff, 233ff; see also Christman 2004 op. cit.; Christman & Anderson, op. cit., p. 3). While not using Rawlsian language, Dworkin's commitment to 'ideological neutrality' might be understood in this way too (op. cit., p. 8).

- 20 Dworkin op. cit., p. 43; see also p. 143.
- 21 John Christman, 'Autonomy and personal history', *Canadian Journal of Philosophy*, 21,1 (1991): 1–24, at p. 22.
- 22 Mill op. cit., pp. 113f; see also Robert Young, *Personal Autonomy: Beyond Negative and Positive Liberty* (London: Croom Helm, 1986).
- 23 See, for example, the discussion in Peter Scourfield, 'Reflections on the serious case review of a female adult (JK)', *The Journal of Adult Protection* 12,4 (2010): 16–30.
- 24 See, for example, Law Commission, *Mental Incapacity — Law Commission Report No 231* (London: HMSO, 1995), §3.3–3.6. Capacity and competence are sometimes distinguished — for example, by presenting capacity as a psychological term that comes in degrees and competence as a binary legal standard. I will treat them synonymously in this article.
- 25 This is Kant's position, or at least one contemporary rendering of it (David Velleman, 'A right to self-termination?', *Ethics* 109,3 (1999): 606–628).
- 26 This is the position found in legal statutes (see, for example, Department of Constitutional Affairs, *Mental Capacity Act 2005: Code of Practice* (London: The Stationery Office), section 4.2; also available at: <http://webarchive.nationalarchives.gov.uk/+http://www.dca.gov.uk/legal-policy/mental-capacity/mca-cp.pdf> (last accessed 26/11/13 at 14.52h)) and those writing on decision-making capacity with a view to legal statute (see, for example, Jillian Craigie & Alicia Coram, 'Irrationality, mental capacities and neuroscience' in N.A. Vincent (ed.) *Neuroscience and Legal Responsibility* (Oxford: Oxford University Press, 2013), chapter 4; Jules Holroyd, 'Clarifying Capacity: Values and Reasons' in Radoilska op. cit., chapter 7, esp. p. 147).
- 27 This view dominates in the philosophical literature not directly anchored in specific statutes (see, for example, Benson 1994 op. cit., 653; Christman 2004 op. cit., pp. 148, 151; 2005 op. cit., p. 281; 2009 op. cit., p. 172; Kristinsson op. cit., pp. 257f; Mackenzie op. cit., pp. 519f).
- 28 See, for example, Benson 2005 op. cit., pp. 124f; and Christman & Anderson, op. cit., p. 3.
- 29 Christman & Anderson op. cit., p. 3. Strictly speaking this should say 'no *direct* constraints', since demanding that the decision-making process is, for example, free from manifest inconsistencies is still a constraint 'on the content of a person's choices or the reasons he or she has for them', albeit an indirect constraint.
- 30 Benson 2005 op. cit., p. 125; my emphasis; see also p. 133.
- 31 Consider Benson's characterisation of normative competency accounts as demanding '[. . .] that agents' capabilities of perception, reasoning and motivation *be connected* in the right sorts of way to what is really valuable or reasonable for them' (2005 op. cit., p. 134; my emphasis).
- 32 This seems to be Benson's own view since 1994 (see 1994 op. cit.; see also 2005 op. cit., p. 125).
- 33 Benson 2005 op. cit., pp. 133f.
- 34 *Ibid.*, pp. 133, 135.
- 35 *Ibid.*, p. 125; see also pp. 136, 137. Benson calls indirectly substantive accounts 'weak substantive' and directly substantive accounts 'strong substantive' conceptions. My phrasing in terms of directness/indirectness is less ambiguous, since a directly substantive account can be strong in one sense (restricting the content of people's choices), but weak in another (involving a single, relatively permissive value that is widely accepted).
- 36 John Christman, 'Saving positive freedom', *Political Theory*, 33, 1 (2005): 79–88, here 84f.
- 37 Christman 1991 op. cit., p. 13; emphasis added.
- 38 Jillian Craigie, 'Competence, practical rationality and what a patient values', *Bioethics* 25,6 (2011): 326–33, at p. 329; emphasis added.
- 39 Craigie & Coram op. cit., pp. 88f; emphasis added; see also Banner op. cit., p. 1040.
- 40 Sarah Buss, 'Personal autonomy' in E.N. Zalta (ed.) *The Stanford Encyclopedia of Philosophy* (Summer 2013 Edition), URL = <http://plato.stanford.edu/archives/sum2013/entries/personal-autonomy/>. (Last accessed 4 November 2013 at 15.26h.), sect. 2; see also Craigie & Coram op. cit., pp. 88f.
- 41 See also Harold Y. Vanderpool, 'The competency of definitions of competency', in M.A. Gardell & E.E. Shelp (eds), *Competency: A Study in Informal Competency Determinations in Primary Care* (Dordrecht: Kluwer Academic Publishers, 1991), pp. 197–210, at p. 203.
- 42 See, for example, Christman 1991 op. cit.; 2009 op. cit.
- 43 See, notably, Geneva Richardson, 'Mental capacity at the margins: The interface between two acts', *Medical Law Review* 18 (2010): 56–77, esp. pp. 68–72; see also Banner op. cit.; and Banner & Szumkler op. cit., p. 381.
- 44 Available at <http://www.legislation.gov.uk/ukpga/1983/20/contents> (last accessed 26 November 2013 at 12.51h). Similar dual regimes exist in other liberal jurisdictions (for a sample overview see

- <http://autonomy.essex.ac.uk/international-legal-comparison-mhamca> (last accessed 26 November 2013 at 12.52h)).
- 45 Similarly, public health considerations may override a competent person's right to decide about his or her treatment, but this is less relevant here, since such cases are not (typically) about paternalistic interventions to prevent self-harm.
- 46 See Andrea Sangiovanni, 'Justice and the priority of politics to morality', *Journal of Political Philosophy* 16.2 (2008): 137–64.
- 47 Rawls initially focuses his 'model case of an overlapping consensus' on a religious view, two comprehensive liberal doctrines (Mill's and Kant's), and a non-systematic ('pluralist') worldview, but later brings in classical Utilitarianism too (Rawls 1993/1996 op. cit., pp. 145–6, 169–71).
- 48 See Freyenhagen op. cit., section 2.
- 49 For one example of this recent trend, see <http://www.fca.org.uk/news/ps14-16-detailed-rules-on-the-price-cap-on-high-cost-short-term-credit> (last accessed 6 February 2015 at 16.22h).
- 50 John Rawls, *Justice as Fairness: A Restatement*, ed. E. Kelly (Cambridge, MA: Harvard University Press, 2001), pp. 18f; see also Rawls 1993/1996 op. cit., pp. xlv, 19.
- 51 Rawls 2001 op. cit., §45.
- 52 *Grundgesetz für die Bundesrepublik Deutschland* [1949], Art. 9(3); as per its official translation available at: <https://www.btg-bestellservice.de/pdf/80201000.pdf> (last accessed 6 February 2015 at 16.22h).
- 53 This line of argument takes inspiration from Rawls's own comments about workplace democracy in Rawls 2001 op. cit., §52.3.
- 54 For comments and criticisms, I would like to thank Viv Ashley, Jill Craigie, Bob Goodin, Timo Jütten, Wayne Martin, Tom O'Shea and Jörg Schaub as well as anonymous referees for this journal. Support for the research presented here was provided by the UK Arts and Humanities Research Council (grant AH/H001301/1), and is hereby gratefully acknowledged.