

In Defense of Moderate Neutralism

Michael Pendlebury

In this essay I discuss the *possibility* and *desirability* of a politically neutral state, using George Sher's recent book on neutrality as a useful foil for much of what I want to say.¹ My thoughts are meant to apply quite generally (although often indirectly) to constitutional democracies around the turn of the millennium.²

Throughout the essay I will use the term "political neutrality" (which I will often abbreviate to "neutrality" simpliciter) for neutrality between competing conceptions of the good that are accepted in a given society. I will be concerned exclusively with *state* neutrality, which I take to include neutrality of public policy, government and administrative practices, low-level institutions, structures and procedures, specific laws and executive acts, and so on. I will refer to these indifferently as "actions or instruments of the state." The possibility and desirability of neutrality of general principles of justice such as the two principles of Rawls's "Justice as Fairness" are somewhat less contentious, and this is perhaps also true of neutrality of what Rawls describes as "the basic structure of society."³ Although there is considerable room for disagreement here, I shall assume for the sake of simplicity that neutrality is both possible and desirable at these levels.⁴ The position I shall advocate is that state neutrality is not fully achievable (part 1), but that it is nonetheless a qualified ideal for which we should strive, subject to various constraints (part 2).⁵

1. The Possibility of a Politically Neutral State

It is widely accepted that the results of state action will often favor some conceptions of the good over others. In the face of this Sher attempts to save the claim that state neutrality is possible by adopting the Larmore-Kymlicka distinction between (as Sher puts it) *justification neutrality* and *effect neutrality*.⁶ This allows him to argue (in line with Larmore and others) that political neutrality is best understood as justification neutrality and that justification neutrality, unlike effect neutrality, is indeed possible.⁷ My main aim in part 1 is to cast doubt on this strategy by questioning Sher's case for the possibility of a justification-neutral state and arguing for the conditional thesis that *if* an effect-neutral state is impossible, *then* in all probability a justification-neutral state is impossible too.

It is tempting to define an action or instrument of the state as justification neutral if and only if its justification is neutral between competing conceptions of the good.⁸ However, this is not very useful for two important reasons.

First, it does not tell us what it is for a justification to *be* neutral between different conceptions of the good. Second, as Sher points out, “any law, policy, or governmental structure, and hence also any decision to implement it, can be justified in many ways” (BN, 23). Given the complexity of governments, the different motives of the numerous players involved,⁹ the inevitable deals and compromises, and the number of individual decisions and actions on which any state action supervenes, it is not clear that it is always possible to identify “*the* justification” in terms of which such an action should be evaluated for neutrality (cf. BN, 23–24).

With regard to the first point, Sher simply assumes that a neutral justification is one that does not involve normative premises to which any conception of the good is committed. He does not say this explicitly, but it is clearly presupposed by much of what he says about neutrality. This applies especially to his discussion of the notion of a conception of the good (see BN, 37–43), in which he feels constrained to exclude the values that he regards as neutral from all particular conceptions of the good, thus ruling out the possibility of common ground between them. But if this were correct, then Rawls’s talk of an overlapping consensus between different “comprehensive doctrines” and the conceptions of the good that they involve would make no sense, since they would not overlap (or would not overlap in their normative commitments).¹⁰

In order to make sense of the way in which Rawls and others use the notion, we cannot understand a conception of the good merely as a contentious belief, judgment, or claim about value, as Sher sometimes suggests (see, e.g., BN, 37, 42, 43), but must construe it as something like a theory (or folk theory) of value. In other words, it is a systematic set of beliefs that is centrally concerned with value and includes in particular beliefs about the nature of the good life for human beings and/or the members of a particular group, including those who hold the beliefs in question. Moreover, there is no good reason to reject the possibility of some areas of agreement between competing conceptions of the good. Allowing for this, Sher could secure the result he is after by defining a justification as neutral between competing conceptions of the good if it involves no normative premises that are *proper* to any of them or whose negations are proper to any of them. To apply this, we could define a proposition as proper to a conception of the good if and only if it is implied by that conception but not by at least one competing conception of the good. The net effect is that only premises that do not belong to undisputed common ground or unoccupied territory can compromise the neutrality of a justification.¹¹ This does reasonable justice to the stereotypically liberal conception of neutrality that Sher wishes to capture.¹² I will therefore refer to it as “the liberal account,” even though there are no doubt liberals who would want to reject it.

I contend that the liberal account of neutrality is unsatisfactory because it fails to capture the fundamental idea that a justification is not neutral between competing conceptions of the good unless it is *equally acceptable* from the perspectives of *all* of them.¹³ And the importance of this idea cannot be overemphasized. Advocates of neutrality often endorse it primarily as an ideal of evident fair-handedness in plural societies in which several compet-

ing conceptions of the good are present,¹⁴ and any alleged form of neutrality that does not meet the requirement of equal (rational) acceptability from the perspective of all competing conceptions of the good cannot hope to satisfy this ideal. This thought is in line with one central feature of Rawls's ideal of public reason, namely, his principle of reciprocity. In terms of this principle, public reasons must be reasons that those affected by the matter under consideration "can not only understand . . . but reasons that we might reasonably expect that they as free and equal might reasonably also accept."¹⁵ Of course public reason as understood by Rawls applies only to "constitutional essentials and matters of basic justice"¹⁶ and not, for example, to everyday legislation. However, this merely confirms that Rawls does not endorse state neutrality (either explicitly or implicitly) in *Political Liberalism* (cf. note 3).

What I have called the liberal account of a neutral justification (as one which merely excludes premises with respect to which competing conceptions of the good differ) fails to meet the requirement of equal rational acceptability from the perspectives of all relevant conceptions of the good. Consider the case of a justification that essentially involves a normative proposition that is not proper to or rejected by any competing conceptions of the good but is more *congenial* to some than to others, perhaps in the sense that it is more likely to be accepted by those who rationally endorse the former conceptions or more likely to be rejected by those who rationally endorse the latter. Such a justification is clearly not equally acceptable from the perspectives of all relevant conceptions of the good, and it should therefore not qualify as neutral.

Furthermore, a justification may be nonneutral not because any of its premises are nonneutral, but because it does not have the resources to address possible objections from the perspective of a particular conception of the good in a neutral way. Consider an argument to the effect that a certain practice should be continued because it is deeply entrenched (which is in itself a neutral consideration, and often one that is not at all negligible). Such an argument should clearly not be counted as neutral if the practice significantly favors some conceptions of the good over others, as it well might if the distribution of power among citizens has been unequal over an extended period. This is a point that Sher should appreciate given his explicit commitment to justification coherentism (see *BN*, 145–47), in terms of which the overall strength of an argument depends significantly on systemic factors, including its capacity to respond effectively to countervailing arguments.¹⁷

This brings me to the second problem that I mentioned earlier, namely, that there may be no particular, identifiable justification relative to which the neutrality of an action or instrument of the state should be evaluated. Sher's approach to this problem can be characterized in general terms as an attempt to do justice to the idea that a principle, constitutional framework, or action or instrument of the state should be counted as justification neutral if it has *at least one* neutral justification that is *adequate* (even if this justification is not publicly accepted or is not entertained by the parties involved).¹⁸ To illustrate: in terms of this general strategy we must count Rawls's two basic principles of justice as justification neutral if we think that his veil-of-ignorance argument provides an adequate neutral justification of those principles—and this

holds even if other possible justifications of the principles are plainly non-neutral. The approach is simple and attractive and avoids setting too high a standard for justification neutrality. I therefore wish to endorse it, especially as it leaves open the possibility of a measure's qualifying as neutral when most of the public discussion and debate surrounding it is plainly partial, which is surely inevitable in the real world.¹⁹

The upshot of all this is that an action or instrument of the state is to be counted as justification neutral if and only if it has at least one adequate justification that is equally acceptable from the perspectives of all competing conceptions of the good. Furthermore, judgments about the rational acceptability of a justification from a given perspective must be based not only on the contents of its premises but also on systemic factors such as those I have outlined. It is obvious that many individual measures can be justification neutral in terms of this account. Consider the law requiring motorists to drive on the left or right (as the case may be) or improvements in the efficiency with which the Income Tax Act is administered. There is, however, good reason to doubt the possibility of a state's being justification neutral *in general*.

To begin with, many significant (and possibly desirable) goals of public policy, which often play an important role in justifications of specific measures, are simply not neutral. Here I am directly at odds with Sher, whose case for the possibility of state neutrality rests heavily on examples of common policy goals that he claims to be neutral on the ground that "virtually everyone agrees that . . . [they] are good things" (*BN*, 38). These goals, which many liberals would also count as neutral, include equality, personal autonomy, freedom, safety and security, health, and economic growth (see *BN*, 1–2, 25, 33, 38). However, it is clear on reflection that none of these is in fact neutral in all societies, for in each case (as I shall argue) there are actual conceptions of the good that are hostile to the goal in question or would have good reason to be hostile to the most likely instantiations of that goal.

It is easy enough to come up with religious conceptions of the good (both orthodox and fundamentalist) that are opposed to equality and reject autonomy and freedom for some categories of people (e.g., women). One does not have to look much further to discover conceptions of the good that are at odds with the goal of safety and security insofar as they see war or conflict as necessary for the identity and well-being of members of the groups or cultures to which they are meant to apply. I heard such a position endorsed on South African television by an extreme right-wing Afrikaner leader in the 1980s.²⁰ It is no doubt because of such cases that Rawls seeks an overlapping consensus only between what he counts as *reasonable* comprehensive doctrines (and the conceptions of the good they involve).²¹

But even reasonable conceptions of the good can differ considerably over both health and economic growth as policy goals. There are many incompatible programs that could be viewed as ways of promoting public health, and any one of them could be more favorable to some conceptions of the good than to others. For example, state regulation and control of health care professionals, drugs and medications, medical insurance, or individual activities (like smoking and sex) that could affect the health of the agents or others around them might be much less attractive to conceptions of the good that

place a premium on individual freedom than to those that greatly value security and predictability. Again, the state's investing in primary health care rather than expensive, high-tech life maintenance systems might be much more congenial to aboriginal conceptions of the good than to some contemporary Western conceptions.

The value of economic growth—as measured by accounting systems approved by the United Nations, the International Monetary Fund, the World Bank, and the global financial sector—is dubious from several perspectives, especially as it gives weight only to monetary transactions, regardless of the worth of the commodities or services involved. The repair of a broken window counts for as much as a night at the opera if they cost the same, and both qualify as infinitely more valuable than a lazy afternoon with a friend or a peasant home built by a family for itself entirely from material found in the vicinity. There are circumstances in which economic growth is an essential goal of responsible public policy (as in most developing countries), but this does not gainsay the fact that economic growth can be threatening to reasonable conceptions of the good that place a high value on leisure, production for consumption rather than trade, or forms of community in which personal and economic relations are not strongly differentiated.

At this point someone might protest that I have not paid enough attention to the distinction between justification and effect neutrality (cf. *BN*, 13). If I did, it might be argued, I would see that even though the effects of economic growth may favor some conceptions of the good over others, it is still possible in appropriate circumstances to come up with at least one adequate neutral justification for economic growth as a policy goal, since it is not necessary for an adequate justification for this goal to trade on the fact that it would harm some conceptions of the good or benefit some more than others.

It is easy to dissipate the force of this objection. To begin with, note that my suggestion was *not* that economic growth and the other policy goals I mentioned cannot be given neutral justifications, but that they are not neutral *as policy goals*, and therefore that no justifications of particular measures that are based on them qualify as neutral justifications. And this, incidentally, holds even in terms of what I have called the liberal account of neutrality.

But I want to go much further and challenge the complacent assumption that adequate neutral justifications of these policy goals are always possible. It must be recognized that, whatever reasons may be invoked in any adequate, presumptively neutral justification of, say, economic growth, *other things being equal* those reasons are apt to be less acceptable from the perspective of conceptions of the good that will *predictably* be prejudiced relative to the others by the achievement of this policy goal. For, if rational, those who endorse these conceptions of the good will have to offset their relative losses against the reasons constituting the justification, thereby diminishing the force of these reasons from their point of view. If this is indeed the case, then the justification is not neutral in terms of the equal-acceptability account of neutrality.

By an obvious generalization from this case, actions or instruments of the state that predictably depart from effect neutrality are in general unlikely to pass the test of justification neutrality. This is a straightforward consequence

of the systemic nature of justification taken in conjunction with the preeminent importance of expected outcomes and predictable effects to political evaluation. From the goal-oriented perspective of politics, justification and effect neutrality are often virtually inseparable.

An anonymous reviewer for this journal has proposed the following counterexample to an earlier (much more rigid) version of the above argument:

Suppose that my conception of the good [includes] having fun via sports, and that I pursue it by playing basketball in my yard. In order to combat crime the government requires that homes be fenced, but the only fence I can build cuts my basketball court in half. Does the fact that this policy goes against my pursuit of my conception of the good show that crime control isn't a neutral justification? I can't see how. Once I know the cost to me of crime control via fencing I may be opposed to it, but that doesn't make the case for it nonneutral.

Given a pinch of salt this seems reasonable, especially because the connection between the cost to the protagonist of crime control via fencing and his conception of the good is an accidental matter. To see this, note that it is very unlikely that similar costs would apply to all or most citizens with relevantly similar conceptions of the good. But if by some grand coincidence they did, I would be open to the thought that this departure from effect neutrality need not compromise the justification neutrality of the measure. And this is clearly consistent with my argument as it is presented above.

Suppose, however, that the government could somehow combat crime by banning sporting activity in private yards after 4 P.M. Then the burdens of this policy to citizens with relevant conceptions of the good would be internal to those conceptions, and this could reasonably be taken to compromise the justification neutrality of the ban. This claim does not imply that the desirability of crime control suddenly becomes a nonneutral consideration, which would be absurd. The point, rather, is that the mere existence of a neutral *consideration* that happens to favor an action or instrument of the state is not sufficient to establish that it is justification neutral, for almost any measure could pass this easy test. This was in fact an important part of the case for the equal-acceptability account of justification neutrality, in terms of which neutrality requires an *adequate justification* of the relevant measure that is equally acceptable from the perspectives of all relevant conceptions of the good.

The above reasoning does not, of course, imply that justification and effect neutrality are conceptually inseparable, and they can undoubtedly come apart in the case of particular measures. However, they are so pervasively interlinked in reality that the possibility of a state that is justification neutral *in general* cannot be saved simply by invoking this distinction.²²

Although this casts doubt on the Larmore-Sher argument for the possibility of a neutral state, it does not establish that a state could not achieve neutrality by restricting its actions entirely to measures that are individually neutral, like the rule of driving on a particular side of the road. But such a minimalist state, even if feasible, might not *as a whole* be neutral between dif-

ferent conceptions of the good. For the overall neutrality or nonneutrality of a state is determined as much by what it permits through inaction as by what it does through active intervention, and it is clear that a minimalist state could easily favor conceptions of the good that are endorsed by powerful and well-established sectors of society, as well as conceptions that place a premium on individual freedom. Whatever their merits, criticisms of limited government by radical feminists and advocates for socially marginalized groups provide significant evidence that a minimalist state is unlikely to be neutral.

A much better way of pursuing real neutrality would involve not total state abstinence from nonneutral measures, but a systematic attempt to balance biases in one direction by compensating biases in others. This sort of dynamic quest to approximate a neutral, balanced whole would have to proceed by the kinds of trade-offs between different conceptions of the good that would be rationally acceptable to all parties on the basis of reasonable knowledge *if there were no inequalities of power or influence between them*. It is extremely doubtful that the neutrality that serves as the nominal goal and limit of such a process could ever be achieved completely. But in any event, this *democratic neutrality* (as I am disposed to describe it) is a far cry from the liberal restraint involving a "prohibition on government efforts to promote the good" (BN, 72) with which Sher is clearly concerned.²³

2. The Desirability of a Politically Neutral State

If political neutrality is not fully realizable, then it cannot be an absolutely necessary condition for the legitimacy of the state (unless we are willing to embrace a philosophical anarchism in terms of which no state is ever legitimate). But political neutrality could still be an ideal for which we should strive, or, more guardedly, a qualified ideal for which, other things being equal, we should strive. Let us refer to the more cautious version of this position as *moderate neutralism*.

Given our commitment to certain fundamental political values, there is, it seems to me, a strong presumption in favor of moderate neutralism. I am thinking in particular of the values of government in the interests of the governed, equality, state impartiality between citizens, and government by rational consent. Since an individual's conception of the good is normally crucial to her interests, departures from neutrality provide *prima facie* evidence for the view that equality and impartiality have not been fully respected, and they also subtract from the reasons that citizens whose conceptions of the good have been prejudiced may have for consenting to government by the state in question. These important considerations support only moderate neutralism because they are subject to the possibility of defeat by additional factors. This opens up the question of the general conditions under which departures from neutrality are legitimate. I will return to this issue after responding to Sher's attack against neutralism.

Apart from insisting that "any neutral state must needlessly cut its citizens off from important goods" (BN, 3), Sher makes his case against the value of neutrality by, as he puts it, "criticising each main argument for neutralism" (BN, 15). He claims that

[b]roadly speaking, neutralism can be defended in three main ways. It can be argued that nonneutral governmental decisions (1) violate the autonomy of citizens, or (2) pose unacceptable risks of oppression, instability or error, or (3) rest on value-premises that cannot be rationally defended [because knowledge of the good is impossible].²⁴ (BN, 15)

However, the first of these arguments is

ambiguous, for it can be taken to appeal either to the great *value* of living autonomously or to a stringent obligation to *respect* autonomy. (BN, 15–16)

This yields a taxonomy of four lines of argument in favor of neutralism, and Sher devotes a chapter to each of them. Although he does not put it thus, in each case his criticism amounts to the charge that the argument does not add up to a demonstrative proof of what I will call *strict neutralism*, namely, a commitment to an absolute and uncompromising prohibition of nonneutral state action.

In chapter 3, for example, Sher considers the argument that he dubs “*the appeal to the value of autonomy*” (BN, 16), which

asserts that each person has a preeminent interest in living an autonomous life and, hence, that government can promote (the most) value by not acting on any particular conception of the good. (BN, 44)

He criticizes this argument on two main counts: first, that it is not at all clear that the state *always* serves the interests of its citizens best by promoting or respecting autonomy rather than other values (see BN, 56–60), and second, that nonneutral state actions need not *always* involve a sacrifice of autonomy (see BN, 61–71). Both points are well taken. However, nothing that Sher says in support of them casts doubt on the view that *other things being equal* the state best serves the interests of its citizens by promoting or respecting autonomy and that this yields a defeasible presumption in favor of state neutrality. Thus his criticisms of the appeal to the value of autonomy carry the required weight only if the argument is meant to be understood as a proof of strict neutralism.

Setting aside the details of chapters 3–6, I think that Sher is absolutely right to think that none of the four lines of argument he discusses amounts to a proof of strict neutralism. Departures from state neutrality *can* sometimes be consistent with an appropriate recognition of the value of autonomy, with a reasonable principle of respect for autonomy, with the need to protect the state from “the risks of oppression, instability and error” (BN, 139), and with an adequate epistemology of value. This does not, however, amount to any sort of case against moderate neutralism.

In fact the four main considerations with which Sher is concerned in chapters 3–6 all add to the credibility of moderate neutralism. For, other things being equal, the value of autonomy, a reasonable principle of respect for autonomy, the need to protect the state from “the risks of oppression, insta-

bility and error," and an adequate epistemology of value all provide reasons, albeit defeasible reasons, for state neutrality.²⁵ At times Sher appears to recognize this, as when he allows that

there are many reasons to acknowledge a strong presumption for allowing citizens to make their own decisions and exercise control over their own lives (*BN*, 104);

that

[w]e also worry, unfortunately with justification, that by tolerating departures from official neutrality, we risk allowing the state's coercive apparatus to be captured by fanatics, bullies, or worse (*BN*, 2);

and that

[n]eutralism also draws support from our uncertainty about where our deepest values lie and which ways of living really are best. (*BN*, 2)

Acknowledgments like these suggest that Sher's position is consistent with moderate neutralism.

However, it is not clear that Sher is willing to accept the key consequence of moderate neutralism, namely, that the nonneutrality of an action or instrument of the state is always a ground to challenge it, even if it does not ultimately defeat it.²⁶ More expansively, the nonneutrality of a measure establishes a presumptive burden on those who support it to show that it does not unreasonably violate the principles of impartiality, equality, and consent; that it does not undermine personal autonomy or the conditions required for the safety and security of the state without adequate justification; and that it is broadly consistent with reasonable beliefs about the nature of the good.

I imagine Sher responding that, although he is willing to grant that as a matter of fact neutrality is often, or even normally, desirable, he can avoid the burdens and presumptions of moderate neutralism by insisting that an action or instrument of the state should always be evaluated strictly "on its merits" (*BN*, 71). This presumably involves the application of "our normal standards of justification" (*BN*, 131) without "an additional layer of argument" (*BN*, 247), for

[t]here is . . . [he claims] nothing to be gained, and much to be lost, by imposing artificial limits on the reasons . . . [that may be considered by political decision-makers]. (*BN*, 248)

Such thoughts undoubtedly carry weight against the unbridgeable constraints of strict neutralism. They are, however, impotent against moderate neutralism, unless they are taken to require that public political reasoning should bypass all *prima facie* principles and presumptions in favor of exceptionless first principles and ultimate moral intuitions. Perhaps this is not quite

what Sher intends, but it is still worth mentioning two important reasons for rejecting the proposal.

The first is that we are simply not up to the task, even in the most theoretical of contexts. Our serious moral reasoning depends heavily on judgment and has a particularist dimension that cannot be eliminated;²⁷ by the standards of mathematics and natural science it is always extremely messy; and it is permanently subject to easy revision. Furthermore, the reasons that morally sensitive human beings offer in support of difficult evaluative judgments correspond at best to *prima facie* generalizations rather than to exceptionless universal laws, and these reasons are often acceptable in terms of our normal standards of justification. Methodological fanatics, like Kant, who insist that any reason must be absolutely and directly universalizable, will inevitably endorse judgments that the rest of us cannot stomach because of their insensitivity to the subtleties, complexities, and ironies of the human condition.²⁸

The second reason for rejecting the proposal that actions and instruments of the state should be evaluated without reference to defeasible presumptions like that in favor of neutrality is that it is politically dangerous. For even if rationally adequate evaluations of this nature are in principle possible, in politically challenging cases they would have to be enormously complex in order to accommodate all possibly relevant considerations. In fact they would have to be so complex that they would be beyond the grasp of the vast majority of citizens. I will not protest that this would be at odds with the *prima facie* principle of consent, since even though this is so, it begs the question. My concern, rather, is that state actions justified in terms of evaluations that citizens cannot understand may turn out to be *unjustified* (despite the supposed skill and good will of the expert evaluators), and they may therefore cause substantial harm to the state and its citizens. In the penultimate sentence of his book, Sher insists that “[i]n the end, we have no choice but to trust the rational faculties of those who will decide” (*BN*, 248). It should, however, be clear from the history of the twentieth century that this is far from adequate as an alternative to institutional structures and public political cultures that tend to promote the satisfaction of appropriate political principles, including a *prima facie* principle of neutrality.

In any case, the burdens and presumptions of moderate neutralism in no way prevent the evaluation of state action “on its own merits,” and they are certainly consistent with our normal standards of justification. Of course *prima facie* presumptions sometimes add extra layers to our moral reasoning. However, this is enabling rather than prohibitive, for it blocks no legitimate considerations but permits simpler modes of argument that are in line with our intellectual powers and the level of precision appropriate to moral and political thought²⁹ and that are also accessible to and assessable by a far wider range of ordinary citizens.

Moderate neutralism survives Sher’s arguments even if we push them much further than he intended. But even his case against strict neutralism is defective. The fact that strict neutralism is not decisively proved by certain arguments does not establish that it is a false ideal. Furthermore, what Sher succeeds in showing is that, *taken alone*, each of the many arguments he con-

siders leaves some space for departures from state neutrality. But the non-neutral measures permitted by one of these arguments could easily be ruled out by another, and Sher does not show that *taken together* they still leave space for nonneutral measures. It is nonetheless clear that the case for strict neutralism is not compelling. Furthermore, anyone with a deep commitment to values like equality, personal autonomy, and freedom should (at least) be suspicious of strict neutralism, since (as I argued in part 1) such values can come into conflict with neutrality.

However, it is unnecessary to pursue this issue, as the liberal neutralists whom Sher is targeting are for the most part committed only to moderate neutralism. This is quite explicit in the case of Larmore, who allows for the "abridgment of neutrality" and for "tradeoffs" between neutrality and other liberal goals.³⁰ In most other cases it is merely implicit, showing up in a variety of ways. These include a flexible attitude toward the question of what state actions are permissible, a commitment to specific nonneutral values, like equality and autonomy, or the hint that the state is not obliged to respect unreasonable conceptions of the good. One liberal who is in his own terms committed to strict neutralism is Ackerman, who takes his principle of discursive neutrality as methodologically fundamental. But Ackerman's conception of neutrality is so latitudinarian that it permits all sorts of reasoning that I would count as plainly nonneutral (see note 13). Moreover, although he is formally a strict neutralist, it is not clear that he would wish to rule out any form of reasoning that Sher would permit. Let us therefore return to the framework of moderate neutralism and take up the important question of the general conditions under which departures from state neutrality are warranted.³¹

By a natural adaptation of Rawls's "difference principle,"³² it might be suggested that departures from neutrality are warranted to the extent that they improve the position of the worst-off citizens *in terms of their own conception of the good*. Although its egalitarian spirit is attractive, this suggestion is not in the end very helpful, in part because it is far too abstract and general. It is also too permissive in some respects and too restrictive in others. Let me illustrate both points in relation to circumstances in which a conception of the good held by some of the worst-off citizens is unreasonable insofar as it clearly fails to serve their own interests. Perhaps they come from a background of extreme social, economic, and educational deprivation and their thinking about the good is greatly influenced by the manipulations of the leaders of an exploitative religious movement. The adapted difference principle must fail on the ground that it would wrongly permit the state to advance the "good" of these citizens as defined by this false conception of their good and would wrongly prohibit it from all possible nonneutral interventions that are inconsistent with that conception of the good but would in fact bring about significant improvements in their lives.

I believe that we can make a more positive first step toward an account of the general conditions under which departures from neutrality are warranted by going back to the relationship between autonomy and neutrality. I have suggested that the value of and the need to respect autonomy provide *prima facie* reasons for state neutrality. This is because nonneutral state action

always (by definition) prejudices some conception of the good and is therefore apt to diminish the extent to which the parties concerned have control of their own lives.

The conception of autonomy I appeal to here is different from Sher's, in terms of which autonomy is construed merely as "responsiveness to reasons" (BN, 56). In my view this puts too much emphasis on autonomous *decisions* rather than autonomous *persons*. But even with respect to decisions, I believe that Sher's account of autonomy is both too broad and too narrow: too broad because it makes a decision based on reasons autonomous even if the person concerned is generally out of touch with reality and at the mercy of his physical and social environment; too narrow because it makes any impulsive decision heteronomous even if it is freely made by an agent who is fully in charge of her own life. The idea of autonomy as control of one's own life integrates and unites a variety of factors, and the significant absence of any of them is felt as a limit to or constraint on one's autonomy. These factors include, most notably, a minimum of external interferences that could restrict or compromise one's choices; the knowledge and understanding to appreciate the threats and opportunities in one's environment and to recognize the consequences of the alternatives between which one is in a position to choose; the psychological competencies required to make such choices rationally and to carry them through; and the absence of significant domination or subjugation by others, which could undermine one's power to choose without actual interference.³³

Why should we care that nonneutral actions or instruments of the state diminish the autonomy of some citizens, especially when there may well be compensating gains in other values? Why is personal autonomy so important that it deserves special consideration from the state? Sher discusses one possible answer to this question, namely, that "only autonomously chosen activities can have value" (BN, 59), and rightly finds it unsatisfactory. But even the much more modest and attractive position that full-fledged human well-being requires autonomy³⁴ cannot, as a general ethical view, explain why autonomy has special *political* significance. It is not, however, difficult to come up with a political explanation of this, namely, that the legitimacy of the state depends on the rational consent of its citizens and that such rational consent in turn presupposes autonomy.

I am not, of course, suggesting that personal autonomy is required for rational consent to *anything*. It is sometimes possible for people who are not autonomous to give rational consent to some things; for example, a child of seven could rationally consent to an afternoon at the zoo rather than a movie, and the same might also apply to someone who is retarded, insane, or subject to general domination by an individual, group, or institution. However, the conditions of rational consent are differentially responsive to different subject matters, and someone who can give rational consent in one field may not be able to do so in another. (This is implicitly recognized by requirements of legal capacity in the law of contract.) And given the deep and pervasive effects of governments on our lives, the rationality or legitimacy of an agent's consent to government, or to particular government policies or actions, is compromised by limitations in personal autonomy.

Thus the thought that the ideal of personal autonomy supports moderate neutralism remains intact. More to the point, this in turn suggests certain conditions under which departures from neutrality are warranted. First, as a political value autonomy is more fundamental than neutrality, so if the two are in conflict, then *other things being equal*, neutrality should give way to autonomy. The same holds, I claim, for neutrality in relation to equality, impartiality, and basic human rights. Second, given the relevance of autonomy to the principle of consent, nonneutral state action is normally warranted if it is required in order to develop the autonomy of citizens. Third, it is *sometimes* in order for the state to favor conceptions of the good that value autonomy (or equality, impartiality, or human rights) over others, especially those that oppose that value. Such departures from neutrality should, however, be carried out only with great caution and sensitivity, for cultural evolution can be slow, and the alienation and resentment that might flow from heavy-handed nonneutral action could do substantial harm to the state as a whole.

This leads on to the general question of the conditions under which state bias against unreasonable conceptions of the good is warranted. Broadly speaking, and setting aside the possibility of internal inconsistency or incoherence, a conception of the good can be unreasonable in one of two ways. I consider these seriatim.

First, a conception of the good can be unreasonable in relation to those who do *not* endorse it,³⁵ insofar as it is apt to be intolerant of their departures from some of its tenets (even if their behavior is consonant with their own conception of the good and cannot reasonably be said to harm or infringe the rights of others). An example of this would be a religion that not only rules out homosexual activities for its members but sees the existence of homosexual activities in the region (or maybe even the world) as a blight on its members' well-being even if these activities are carried out in secret by unknown consenting adults. Although my earlier caveat about the dangers of insensitivity applies here, too, a reasonable principle of reciprocity and mutual respect suggests that the advocates of such a conception of the good do not have the right to demand state neutrality in relation to its unreasonable aspects.

Second, a conception of the good can be unreasonable in relation to some or all of those who *do* endorse it insofar as some of its main tenets are palpably incompatible with their interests. I am inclined to think that any conception of the good that is committed to values in an afterlife in place of earthly well-being belongs to this class. The same applies to any conception of the good that is committed to fundamental inequalities between sexes, castes, ethnic groups, or genetically constituted estates (where some members of all the groups in question endorse that conception). Whether a conception of the good that is unreasonable to some or all of those who endorse it deserves the same respect from the state as a reasonable conception of the good depends, I suggest, on whether those to whom it is unreasonable are autonomous people. If they are, then they have freely hoisted themselves on their own petards, and the state and the rest of the citizenry should merely sit back and wonder at the marvels of human diversity. But if the internal

victims of an unreasonable conception of the good are not autonomous, as in the example I gave earlier, then it seems to me that the state has a presumptive right to intervene.

To conclude, the ideal of government by consent makes the development of a fully autonomous citizenry a crucial long-term desideratum for public policy, and it is inevitable that this goal will sometimes trump short-term considerations of neutrality. This is, however, consistent with moderate neutralism, as well as with a further, more demanding ideal that we can at best hope to approximate in a very remote future, namely, that of an order in which significant departures from neutrality are never required or warranted.

Appendix on Framework Neutrality

The principles of justice and constitutional structures operative in a given society should perhaps be neutral between *fully reasonable* conceptions of the good accepted in that society, but it is not clear that this thought can be unpacked in a way that begs no important questions about the nature of reasonability. However that may be, considerations like those raised in the above essay suggest that framework neutrality between the *actual* conceptions of the good accepted in a society may be neither possible nor desirable given the contents of some of those conceptions of the good. Consider the following examples with respect to the issue of desirability:

1. A principle of minimal formal equality between the members of a society would not be neutral if it were unacceptable from the perspective of regressive conceptions of the good that continue to attract support in that society. This does not, however, mean that such a principle is not required by justice.
2. Rawls's difference principle surely qualifies as a possible principle of justice despite the fact that it is not neutral between egalitarian liberal conceptions of the good and libertarian conceptions of the good.
3. The Constitution of the Republic of South Africa of 1996 lays down a framework of democratic government and a set of principles, goals, and ideals of justice that are widely (and I think rightly) regarded as admirable or even exemplary. This constitution is not, however, neutral between the main conceptions of the good accepted in contemporary South Africa, and there are significant groupings who are hostile to the substantive equality that it seeks to promote.

The recognition that framework neutrality may not be possible or desirable in some societies does not make a significant difference to the arguments concerning state neutrality set out in the above essay. For these arguments could in principle be reformulated to accommodate this complication, often by replacing the unrestricted idea of *neutrality between competing conceptions of the good* by something like the restricted idea of *neutrality between competing conceptions of the good that are consistent with the relevant framework with respect to the matters at hand*.

I am grateful to the members of the Witwatersrand discussion group on political philosophy for stimulation without which this essay would never have been written; to various people for useful comments, including members of my audiences when I presented much earlier versions of this material at the annual conference of the Philosophical Society of Southern Africa, Bloemfontein, January 1999, and to the Philosophy Department at the University of the Witwatersrand in March 1999, and an anonymous reviewer for this journal; and especially to Mary Tjiattas for acting as an ongoing sounding board. I also owe thanks to the University of the Witwatersrand for a University Council Research Grant for the Human and Social Sciences that gave me the time needed to produce the final version.

Notes

¹ George Sher, *Beyond Neutrality: Perfectionism and Politics* (Cambridge: Cambridge University Press, 1997). This is cited hereafter as *BN*.

² These include, among others, newer democracies in the developing world (such as India and South Africa) and a number of states that were formerly part of the Soviet bloc, as well as wealthy, well-established Western democracies (which all too often set the limits of mainstream political philosophy).

³ See John Rawls, *A Theory of Justice* (Cambridge: Harvard University Press, 1971), 60–65 (on the two principles) and 7–11 (on the basic structure of society). As Sher reads him (*BN*, 31–32), Rawls is committed to both state neutrality and framework neutrality in *A Theory of Justice*, but only to framework neutrality in *Political Liberalism* (New York: Columbia University Press, 1993) (in which Sher refers specifically to page 235).

⁴ I stress, however, that this is nothing more than a simplifying assumption. For some further remarks on neutrality at these levels, see the “Appendix on Framework Neutrality” (which is best read after the essay itself).

⁵ The alternatives to this position that are advanced in the literature include the following:

(i) In *Political Argument* (London: Routledge and Kegan Paul, 1965), Brian Barry endorses the view that a politically neutral state is neither possible nor desirable. More specifically, on 74–79 Barry argues for the impossibility of a politically neutral state within the scope of the assumption that its desirability presupposes its possibility—because “ought” implies “can” (74). On the face of it Barry is much more positive about neutrality in *Justice as Impartiality (A Treatise on Social Justice, vol. 2)* (Oxford: Clarendon Press, 1995). However, it is clear from chapter 6 that he still rejects the possibility and desirability of state neutrality as I have characterized it (see especially 143). The neutrality that Barry endorses in *Justice as Impartiality* is an abstract, second-order procedural neutrality that is not at issue in this essay (although it does strike me as a reasonable requirement of justice).

(ii) Sher endorses Barry’s view that the desirability of state neutrality presupposes its possibility, claiming that “there would be little point in going on” to evaluate its desirability if it were not possible (*BN*, 4). He then goes on to argue at length that it is possible but *not* desirable. The present essay contains numerous criticisms of these arguments.

(iii) In *Patterns of Moral Complexity* (Cambridge: Cambridge University Press, 1987), chapter 3, Charles Larmore (whom we may take as a fairly representative liberal) argues that state neutrality is both possible and desirable. Larmore is not, however, committed to the absolute possibility and desirability of neutrality, and at an abstract, general level he could easily accept my view that state neutrality is not fully realizable and that it is merely a qualified ideal. There are nonetheless major differences between our arguments and the details of our positions. These arise largely from the fact that Larmore is operating within a liberal framework whereas I am not.

⁶ See Larmore, *Patterns of Moral Complexity*, 44, and Will Kymlicka, “Liberal Individualism and Liberal Neutrality,” *Ethics* 99 (1989): 883–905, at 883–84. Larmore describes the dis-

inction as that between *procedural* and *outcome* neutrality and Kymlicka as that between *justificatory* and *consequential* neutrality. Kymlicka presents his version of the distinction as an attempt to make sense of Raz's distinction between "the exclusion of ideals" and "neutral political concern." See Joseph Raz, *The Morality of Freedom* (Oxford: Clarendon Press, 1986), chapter 5.

⁷ It is worth noting here that the procedural neutrality endorsed by Barry in *Justice as Impartiality* (see note 5) does *not* imply justification neutrality.

⁸ As previously indicated, the relevant conceptions of the good are those that are *accepted* in the society concerned. Thus the neutrality or nonneutrality of an action or instrument of the state is relative to that society. It is also worth noting here that when I say that a conception of the good is "accepted" in a given society, this should be understood as shorthand for something like "explicitly or implicitly accepted on a more or less ongoing basis by a not insignificant group of members of that society." Although this is extremely vague, it does the required job of excluding conceptions of the good endorsed only by a few crackpots.

⁹ These include citizens, government officials, interested parties and organizations, and politicians of various political allegiances.

¹⁰ See, for example, Rawls, *Political Liberalism*, Lecture IV. It must of course be recognized that for Rawls the overlapping consensus between reasonable comprehensive doctrines need be only broad enough to provide support for his principles of justice, which he places in the domain of the right, not the good. However, his second principle itself involves a commitment to certain primary goods. Furthermore, at this point in my argument all that matters is that the idea of an overlapping consensus should be *intelligible*. The substantive contents of such a consensus are not at issue.

¹¹ It also turns out that a premise that is rejected (either implicitly or explicitly) by all competing conceptions of the good cannot undermine the neutrality of a justification. This result is a bit odd, but Sher is implicitly committed to it anyway. Furthermore, it is not of any great consequence, since no reasonable justification would depend upon such a premise.

¹² Compare, for example, Larmore, *Patterns of Moral Complexity*, 42–47.

¹³ Analogous strictures also apply to other conceptions of neutrality that are at least as latitudinarian as the liberal account, including Ackerman's stipulation that neutrality merely rules out appeals to the presumed superiority of a particular conception of the good. See, for example, Bruce Ackerman, "Political Liberalisms," *Journal of Philosophy* 91 (1994): 364–86, at 369, in which the position advanced in Ackerman's *Social Justice in the Liberal State* (New Haven: Yale University Press, 1980) is reaffirmed.

¹⁴ See, for example, Larmore, *Patterns of Moral Complexity*, 50–55.

¹⁵ On the idea of public reason see Rawls, *Political Liberalism*, Lecture VI, and especially the "Postscript" to this lecture in James Bohman and William Rehg, eds., *Deliberative Democracy: Essays on Reason and Politics* (Cambridge: MIT Press, 1997), 131–41. The quotation to which this note is appended is from "Postscript," 135.

¹⁶ This phrase appears in numerous places in the works cited in note 15, including, for example, "Postscript," 134.

¹⁷ This is, I believe, something that we should accept even if we reject unrestricted epistemic coherentism. I will assume that it is correct in the remainder of this essay.

¹⁸ What Sher actually says is that it should have at least one neutral justification that "contains no implausible premises or obvious fallacies, and . . . provides a justification of reasonable strength" (BN, 26), which is best understood as an account of justificational adequacy. I don't, however, consider the details of this account very enlightening, especially as it fails to draw out the need for an adequate justification to have the capacity to respond satisfactorily to countervailing argument. (This is not to suggest that it should be *stronger* than any possible countervailing argument, which, as Sher points out, would be far too demanding a requirement [BN, 25].) The above criticism does not, however, reflect adversely on the general lines of Sher's approach as I have characterized it.

¹⁹ Although Sher is clearly aware of this benefit of his strategy, he seems to lose sight of it when arguing that justification neutrality prohibits governments, legislators, "and

probably also [their] aides and advisors" from "acting for nonneutral reasons" (BN, 29–30). Even on its own such a prohibition would render state neutrality well-nigh impossible.

- ²⁰ If memory serves me right, the person in question was the late Jaap Marais, leader of the Herstigte Nasionale Party at the time.
- ²¹ See, for example, Rawls, *Political Liberalism*, 134.
- ²² The intimate connections between justification and effect neutrality also make it unnecessary to distinguish between the two in most contexts, and I will generally ignore the distinction in the rest of this essay (as Sher ignores it in most of his book, except when it appears to help him out of a tight corner).
- ²³ In part 2 I try to frame my discussion of the desirability of state neutrality in general terms, ignoring the differences between different conceptions of neutrality, such as the two mentioned in the paragraph to which this note is attached. I do not believe that this simplifying strategy results in any insuperable problems.
- ²⁴ Elsewhere Sher glosses the third supposed way of defending neutralism as a "(broadly) epistemological class of arguments . . . [in terms of which] the reason the state must . . . [be neutral] is simply that no one can know, or know well enough, which ways of living are good" (BN, 17)
- ²⁵ Some considerations that support this appear later. The point is perhaps least obvious with respect to the epistemology of value, but (drawing on Rawls's remarks on "the Burdens of Judgment": see Rawls, *Political Liberalism*, 54–58) the fact that reasonable conceptions of the good differ over values that are not neutral between them is enough to establish that full knowledge is not to be had with respect to these values. Thus the state cannot easily be justified in endorsing some of them in preference to others. Add the (quite reasonable) suggestion that any given conception of the good is to be presumed reasonable unless it can be shown to be unreasonable with respect to the issues at hand, and the nonneutrality of a given value becomes a *reason* why the state should not pursue it.
- ²⁶ That Sher is unwilling to accept this is suggested by the first paragraph of his final chapter, in which he moves directly from his rejection of strict neutralism and his earlier defense of a particular perfectionist account of "the best lives" to "[t]he obvious conclusion . . . that governments . . . often have ample reason to promote such lives" (BN, 245).
- ²⁷ See, for example, Larmore, *Patterns of Moral Complexity*, chap. 1.
- ²⁸ This is not of course intended to gainsay the immense importance, value, and interest of Kant's ethics. Nor is it to reject the principle of universalizability, which is consistent with various modes of moral reasoning, including repeated appeals to prima facie generalizations until we reach acceptable principles to which we do not consider the relevant case to be an exception.
- ²⁹ "We must be content . . . in speaking of such subjects . . . to indicate the truth roughly and in outline . . . ; for it is the mark of an educated man to look for precision in each class of things just so far as the nature of the subject admits" (Aristotle, *Nicomachean Ethics*, 1094b, trans. David Ross, London: The World's Classics [Oxford University Press], 1954, 3).
- ³⁰ Larmore, *Patterns of Moral Complexity*, 68.
- ³¹ In doing this I will not consider the possibility of nonneutral state actions' being warranted for specifically perfectionist reasons, in part because Sher gives no positive arguments for such a warrant (cf. note 22).
- ³² See Rawls, *A Theory of Justice*, 75–78.
- ³³ Here I draw on my article "Personal Autonomy, Democracy and the New Constitution," *South African Journal on Human Rights* 11 (1995): 485–94.
- ³⁴ For a perspective on human well-being in terms of which autonomy is required for well-being without being the basis of other goods involved in well-being, see my article "Against the Careerist Conception of Well-Being," *Philosophical Forum* 31 (2000): 1–10.
- ³⁵ I assume that this may include members of a group that is predominantly committed to the relevant conception of the good, including children and adult members of the group who are unable to embrace it in their hearts.