

Perfectionism, Economic (Dis)incentives and Political Coercion

Oran Moked, Columbia University
(Draft 3.28.2008)

May a government attempt to improve the lives of its citizens by promoting the activities it deems valuable and discouraging those it disvalues? May it engage in such a practice even when doing so is not a requirement of justice in some strict sense, and even when the judgments of value and disvalue in question are likely to be subject to controversy among its citizens? These questions have long stood at the center of debates between political perfectionists and political neutralists. In what follows I want to address a prominent cluster of arguments against political perfectionism—namely, arguments that focus on the coercive dimensions of state action. My main claim is simple: whatever concerns we might have about coercion, arguments from coercion fall far short of supporting a thoroughgoing rejection of perfectionism, for the reason that perfectionist policies need not be coercive. The main body of the paper responds, however, to some neutralist challenges to this last assertion.

§1 Anti-perfectionist Arguments from Coercion

To situate my concerns within the terrain of the neutrality/perfectionism debate, let me begin with what will necessarily be a rough sketch of the debate's main contours.

Perfectionist *moral* theories are theories that ground moral reasoning at least partly in some objective account of what is good or valuable for human beings. Such accounts may be based on a conception of human nature, although that need not be the case: they may simply rely on the identification of certain human activities, capacities, attributes or pursuits as valuable, without commitment to a more robust theory of human nature or essence.¹ Perfectionist *political* theories simply extend perfectionist moral reasoning to the sphere of politics, maintaining that the entire expanse of moral considerations—those rooted in objectivist accounts of value included—may, indeed should, guide all action, whether political or nonpolitical.

What anti-perfectionists, or neutralists (I will use the terms interchangeably), typically claim, by contrast, is that certain kinds of moral consideration ought not to guide, or ought to be excluded from, the justification of certain types of action. Anti-perfectionism does not, however, require that *all* moral considerations be excluded from the justification of the problematic types of action; nor does it require that the problematic moral considerations be excluded from the justification of *all* action. Any anti-perfectionist argument must therefore begin by answering two basic sets of questions: (1) To *which class of considerations* does the exclusionary restriction apply, and why to that class alone? (2) To *whose actions*, or to *which sphere(s) of action*, does the restriction apply, and why only to those?

The typical neutralist answer to the first question is, roughly, that the restriction applies to moral considerations rooted in *conceptions of the good*—in the thick or comprehensive accounts of value provided by perfectionist moral theories—but not to the rest of morality. The typical answer to the second question is that the restriction applies to *state* or *government* action (or to

¹ This corresponds to Hurka's (1993) distinction between, respectively, narrow and broad perfectionism.

political action), but not to individuals acting either privately or nonpolitically. In other words, doctrines of neutrality challenge the move from moral to political perfectionism, demanding that states or governments should not appeal to perfectionist conceptions of value when justifying their policies.²

Some of the difficulties facing neutralist doctrines are connected to the first of our two sets of questions. In particular, there is the familiar worry that there might be no defensible way of dividing the realm of moral considerations into two parts, such that one part may serve as the justificatory basis of political or state action while the other may not. It has been argued, for example, that even the likeliest candidates for membership in the part of morality neutralists deem appropriate for political reasoning—e.g., requirements of justice—cannot but rely for their force and content on the *other* part of morality, on particular conceptions of the good, and that doctrines of neutrality are therefore doomed to incoherence.³

These, however, are not the worries I wish to discuss here. (Let me just note that I do not find them as obviously damaging to the neutralist case as some have taken them to be.)⁴ Instead, I want to address our second set of questions: to whose actions, or to which sphere(s) of action, should anti-perfectionist restrictions apply—and why? Why is it political action, or state action, that merits special concern? If we have compelling reasons to encourage certain conceptions of the good and discourage others, why may we not do so politically, through the state, as well as nonpolitically? Unless anti-perfectionists can argue for some morally relevant distinction between political and nonpolitical action—unless they can drive a wedge between moral perfectionism and political perfectionism—it remains unclear why conceptions of the good should be excluded from political reasoning.

Not all of the by now familiar arguments for state neutrality meet this very basic requirement: arguments from moral skepticism, for example, do not.⁵ Much better at identifying

² I largely sidestep here several key issues: (1) Among what is the state to be neutral: Conceptions of the good in general? Controversial conceptions of the good? Reasonable conceptions of the good? Etc. (2) With respect to what is the state to be neutral: The justification of its policies? Their aims? Their effects? (In what follows I will largely speak in terms of neutrality of justification.) (3) Whose actions are to be subject to neutralist restrictions: The actions of public officials? Of all citizens to the extent that they are acting politically (e.g., as voters)? (For simplification, I will use the term ‘political action’ in the restrictive sense, as referring only to the actions of governments, or states, and their official representatives. I will therefore use the terms ‘political action’, ‘state action’, and ‘government action’ interchangeably.)

³ These worries have been the main motivation for attempts to provide a ‘neutral’ justification of principles of neutrality, that is, an anti-perfectionist argument independent of perfectionist value claims. (See, e.g., Ackerman 1980, 11-12; Larmore 1987, Ch. 3.)

⁴ The core anti-perfectionist claim—that there is a class of moral considerations which are appropriate for moral reasoning outside of politics but inappropriate for political reasoning—need not be self-contradictory, even if it is conceded that any piece of moral reasoning must involve or rely on perfectionist considerations. Even if it is true that morality cannot be divided into conceptions of the good on the one hand and moral considerations that do not depend on any conception of the good on the other, this does not imply that the realm of moral considerations cannot *at all* be defensibly divided into two parts, such that it is at least coherent to think that one part may legitimately be applied to political action while another may not. As Jeremy Waldron (1993, 159ff.) has put it, it is not incoherent to be neutral among *some* conceptions of the good without being neutral among *all*; and if this is true, then it is not self-contradictory to offer a ‘non-neutral’, or perfectionist-dependent, justification for principles that retain the basic exclusionary, two-class structure of anti-perfectionist doctrines—so long, of course, as this is done with the understanding that the doctrine thereby becomes more modest in its aspirations.

⁵ Unless, that is, they are accompanied by other arguments, such as the arguments from coercion I discuss below. Whatever else may be said in objection to skeptical arguments for neutrality (e.g., that moral skepticism is false; that even if true, it cannot coherently be invoked to support any moral doctrine; etc.), taken by themselves they fail to identify any relevant difference between political and nonpolitical action. If the argument applied to the former, it

the requisite distinction, however, are arguments that point to the *coercive* character of state action.

Anti-perfectionist arguments from coercion come in a variety of forms. They vary in their conceptions of what it is that we should worry about when we worry about coercion. (Freedom in some thin sense of arbitrary choice, of not being subjected to the will of others, and so forth? Or some thicker ideal of autonomy as reasoned self-direction?) They can also vary in their conceptions of how freedom- and coercion-related considerations function within morality. (Can some loss of freedom be offset by other sorts of gain? Or do considerations of freedom enjoy some more privileged status?) Concerns about coercion are often connected to the idea that governments may act only on reasons that the governed—at least the reasonable among them—share, or could share, or cannot rationally reject. Such concerns are sometimes associated with the Kantian charge that coercing people for reasons they do not or could not share treats them merely as means rather than also as ends; with the Lockean worry that attempts to ‘coerce people into the good’ are self-defeating; with the Millian idea that uncoerced experimentation with various conceptions of the good is instrumental to finding out what is in fact good; etc.⁶

Of course, hardly anyone thinks that all coercive acts are unjustified: some actions ought to be prohibited, and the prohibitions coercively enforced. The apprehension, however, is that to ground such prohibitions in perfectionist reasoning is to expand the use of coercion to matters that ought to remain free of coercive regulation.

But although anti-perfectionists so regularly base their arguments on the coercive aspects of state action, what they typically argue for excluding are not just legal prohibitions grounded in perfectionist reasoning, but *any* kind of political action so grounded. In doing so, they often ignore the fact that perfectionist action need not be the “coercive imposition of a style of life”:

Conferring honours on creative and performing artists, giving grants or loans to people who start community centres, taxing one kind of leisure activity, e.g., hunting, more heavily than others, are all cases in which political action in pursuit of conceptions of the good falls far short of the threatening popular image of imprisoning people who follow their religion, express their views in public, grow long hair, or consume harmless drugs.⁷

It would seem, then, that arguments from coercion do not suffice for a thoroughgoing, morally principled rejection of political perfectionism.⁸ This is underscored by the fact that economic (dis)incentives—subsidies, taxation of specific activities—as well as other noncoercive policies are not marginal, easily negligible components of perfectionist practice, but rather commonplace instruments of perfectionist state policy.

Now, to point to the fact that perfectionist policies need not be coercive is not yet to discredit all coercion-based arguments for state neutrality: it is only to show that coercion-based arguments fail to support *strong* versions of neutralism, according to which *no* government policies may be justified on perfectionist grounds. It does not show that coercion-based

would likewise apply to the latter: it should place a constraint on all moral reasoning, or it should place no constraint at all. (See Waldron 1993, 158-59; Raz 1994, 99.)

⁶ Neutralist arguments that give pride of place to considerations of coercion are found (to mention just a few prominent examples) in Rawls 1971 (328-29) and 1993 (216ff.), Nagel 1987, Kymlicka 1990 (222-23), and probably Larmore 1987 (given a suitable interpretation of Larmore’s ‘norm of rational dialogue’).

⁷ Raz 1986, 161. See also Hurka 1993, 158-60; Sher 1997, Ch. 3; Wall 1998.

⁸ ‘Morally principled’, for there may well be reasons of a more pragmatic nature for limiting the power of governments to promote what they take to be good. (On the limitations of pragmatic arguments, however, see Sher 1997, Ch. 5.)

arguments could not support some weaker principle of neutrality which applied only to coercive state action.⁹ (To reject the weaker principle, one would have to argue that perfectionist coercion *can* sometimes be justified.)

But perhaps—or so some have suggested—even ‘strong’ neutralists could respond to the above perfectionist challenge. To argue that coercion-based arguments support even ‘strong’ neutralism, neutralists would have to resort to one or more of the following strategies: (1) They could question whether the very distinction between coercive and noncoercive state action is plausible. (*The argument from ultimate coercion*). (2) They could argue that even if such a distinction can be made, allegedly noncoercive policies such as economic (dis)incentives *do* in fact coerce people into accepting (or rejecting) conceptions of the good, or into acting (or refraining from acting) in accordance with such conceptions. (*The argument from direct coercion*). (3) They could argue that even if perfectionist policies are not coercive, they are nevertheless *manipulative*, and that the same reasons that weigh against coercive state actions also weigh against manipulative ones. (*The argument from manipulation*). Or perhaps (4) they could claim that however we classify policies such as subsidization or activity-specific taxation, what is objectionable about them is that they induce people to act for the ‘wrong reasons’, and that such influence, whether properly speaking ‘manipulative’ or not, similarly undermines people’s autonomy. (*The argument from inappropriate motivation*.)¹⁰ In what follows I want to respond to these arguments.¹¹

§2 Shifting the Burden of Argument?

Before I do so, however, let me address one possible objection to the general argumentative strategy I employ here, which, to recap, is this: I have said that if we have reason to promote valuable goals, and if doing so by nonpolitical means is acceptable, anti-perfectionists need to explain why doing so by political means is not. To do so, they need to point to some relevant distinction between political and nonpolitical (or state and non-state) action. And if the distinguishing feature to which they point is the state’s coercive character,

⁹ Cf. Sher 1997, 37. Arguing against strong neutralism is nevertheless worthwhile, I think, not only because that, rather than weaker variants, is the sort of doctrine neutralists have usually defended, but also because it is the stronger variant that is truly distinctive, the one whose requirements are not assimilable to those of non-neutralist principles of *toleration*.

¹⁰ Versions of these claims have been presented most thoroughly by Waldron 1989 (to whom I shall devote the most attention), and more recently in Gaus 2003. They are also implied, however, by the rather persistent claim that “all political power is coercive power,” found for example in Rawls 1993, 216.

¹¹ I will not address a fourth significant challenge here, but let me say a few words about it anyway. Anti-perfectionists might also claim, of course, that even when perfectionist state policies neither coerce nor manipulate people into accepting or rejecting conceptions of the good (or the activities associated with them), the state needs to resort to coercive means—primarily taxation—in order to fund its policies; and that such coercive means, though defensible when the state excludes conceptions of the good from its justificatory practices, are objectionable when it does not. One easy response—too easy—would be to say that perfectionist government programs can be funded on a voluntary basis, for example by lottery proceeds. But any attempt to defend a wider range of real-world funding methods would have to acknowledge the presence of coercion, yet insist either that the resultant loss of freedom is somehow necessary for the promotion of an even greater degree of freedom or autonomy, *or* that some measure of freedom can sometimes be sacrificed to promote values other than freedom. One important point to keep in mind, though, is that *what* is being sacrificed when taxes are coercively collected is a good of relatively lesser import—money, or, rather, the freedom to use it as one pleases—and *not* one’s deep convictions or cherished pursuits.

they need to defend their view against the objection that many state actions of the sort perfectionists recommend are in fact noncoercive.

But an anti-perfectionist objection different from the ones just noted might be available. Instead of challenging the move from moral to political perfectionism, neutralists might try to block an earlier move *within* moral perfectionism. That is, neutralists might claim that whatever the differences between political and nonpolitical action, in *neither* case do we have a reason to intervene in order to encourage others to adopt worthy ideals or discourage them from adopting unworthy ones. As Jeremy Waldron puts it:

...the connection between [perfectionist] evaluation and perfectionist intervention is not nearly so straightforward. That an ideal is unworthy provides a person with a reason not to choose it as her ideal, but it is not at all clear that it provides others ... with a reason to discourage her from choosing it.¹²

It is unclear, however, what might justify such a view—unless, that is, it rested on a moral theory that was thoroughly egoistic, at least with respect to perfectionist goals.¹³ Without here arguing against egoistic versions of moral perfectionism, let me just note that apart from what I take to be their utter unattractiveness, they also seem to be at odds with the primary thrust of doctrines of neutrality, which are typically about the relation between moral reasoning in general and moral reasoning in the political sphere (seeking as they do to exclude some moral considerations—appealing as they may be—from politics), and not about the content of basic morality itself.

Let me also add that my initial claim concerning a (defeasible) reason to intervene should not be confused with another, stronger claim—namely, with the claim that the unworthiness (say) of a person's ideal imposes on others a *duty* to discourage her from pursuing it. Even non-egoistic versions of moral perfectionism that made this stronger claim would be vulnerable to a variant of Waldron's objection that substituted 'duty' for 'reason'. But my argument sidesteps this possible complication, for my initial claim invokes only a reason, not a duty, to intervene. And so, the request that anti-perfectionists provide overriding considerations still stands. To the above-noted anti-perfectionist attempts to do so I now turn.

§3 The Argument from Ultimate Coercion

The first anti-perfectionist response to the claim that perfectionist state action need not resort to coercive measures (and thus to the claim that the state's general coercive character fails to provide the requested overriding considerations), is to question whether it is at all meaningful to divide the means available to the state into coercive and noncoercive ones. Because all influence wielded by the state ultimately rests on the state's coercive power, this line of reasoning goes, all state actions are *ultimately* coercive, even when no overt coercion is involved: "The state may not show its guns when it takes [seemingly noncoercive] action. But ultimately it is what it is on account of its guns."¹⁴

How seriously should we take such claims? Not very, I should think. Many state actions and policies (military conscription, mandatory schooling) are coercive in nature; many others

¹² Waldron 1989, 1136-37.

¹³ This would include mixed moral theories combining perfectionist and non-perfectionist considerations, where only the former are egoistic.

¹⁴ Waldron 1989, 1152 (cf. 1139-40, 1143). See also Gaus 2003, 11; and, more implicitly, Rawls 1993, 216.

(the designation of national holidays, the granting of public honors) are not. (That actions of the latter kind sometimes rely on further measures that *are* genuinely coercive does not make them coercive themselves: museum subsidies force no one to visit museums or to adopt the conceptions of value that such visits embody, even when the subsidies are funded by coercively collected taxes.) The idea that the state's considerable coercive power contaminates *all* state action is no more plausible than the Austinian view that all legal rules are commands backed by coercive threats. Like this long-rejected legal-philosophical position, the argument from ultimate coercion blurs distinctions precisely where distinctions ought to be drawn.

This is not to say that we cannot identify a possible source of whatever appeal the argument from ultimate coercion does have. There *are*, after all, cases in which it might indeed be sensible not to treat differently coercive demands and noncoercive recommendations or requests. Requests we would otherwise treat as noncoercive can function coercively when issued by a superior—a supervisor, a teacher, an officer—to a subordinate. Such requests need not be communicated explicitly (and perhaps not even by way of insinuation) as coercive threats in order to function as ones. Arguably, to have a coercive effect such requests need not even be intended as coercive threats, nor interpreted by their recipients as so intended. Instead, it might be enough that the recipients have reason to fear that failure to comply might result in some (perhaps yet unplanned or un contemplated) act of reprisal. Those, for example, who propose that sexual harassment laws indiscriminately prohibit all sexual communication between superiors and subordinates are likely to appeal to precisely this sort of reasoning.

The question, then, is whether perfectionist state policies that are not coercive in any apparent or straightforward way may nevertheless be coercive (and, given further argumentation, objectionable) for the sort of reason described just now. The type of case analogous to the ones discussed in the previous paragraph would be one in which a government issued a recommendation without communicating in any way an intention to penalize non-compliers, and yet citizens had reason to fear that failure to comply would carry penalty after all. But should such clear instances of foul play form the basis of a principled position on the kinds of moral reasoning admissible in politics? Should we not rather consider cases in which the government does not violate its own publicly announced rules—cases in which the state refrains from using any guns it has not previously shown? Are there any principled, coercion-related reasons for rejecting political perfectionism even then? Not in any way that the argument from ultimate coercion can establish.

§4 The Argument from Direct Coercion

To challenge the claim that perfectionist state action need not be coercive, then, one must point to more than just the state's general power to coerce: one must point to some genuinely, directly coercive aspect of seemingly noncoercive policies. Since this argumentative strategy can only proceed piecemeal, one policy at a time, even those who find it promising should not expect it to yield a complete case against political perfectionism. But even partial success would be of consequence if the policies exposed as coercive were central enough instruments of perfectionist legislation.

One such instrument—the one most clearly susceptible to the present type of argument—is the taxation of specific activities that a government wants to discourage. Although activity-

specific taxation falls far short of criminalizing the discouraged activities, its impact on individuals' choices, the argument goes, is nevertheless that of *coercive threats*:

What does it mean to discourage an activity by imposing a tax? It means ... that the state raises the costs that a person must incur if she wants to pursue the activity. ... Now, certainly, altering the costs and payoffs of an activity *looks* like coercion. After all, what a threat does ... is precisely to add an artificial cost to an activity. The threat, 'Your money or your life,' ... adds the cost of death to the option of keeping one's money.¹⁵

(I call this an argument from *direct* coercion, since the suggestion here is that activity-specific taxation directly coerces people into refraining from the taxed activity, and is not merely contaminated by coercion in a more roundabout way.)

Is the challenge reasonable? The argument, as presented in the quote, equates activity-specific taxation and other (dis)incentives with the adding of 'artificial' costs;¹⁶ the adding of 'artificial' costs with threats; and, implicitly, threats with coercion. But although each of the last two moves (in one version or another) has had its defenders in the literature on coercion and threats, by no means are they unquestionable. (Let us grant for now that the first move—the claim that government (dis)incentives add 'artificial' costs to the activities they target—is warranted. I will return to this point in the next section.)

The first thing to note is that not every proposal to add costs to an option can plausibly be said to constitute a *threat* to those who want to pursue the option. [EXAMPLE.] Let us suppose, however, that this difficulty is addressed by the reference to '*artificial*' costs, which implies the existence of some 'natural' or otherwise appropriate baseline of costs and benefits, deviation from which is necessary if an act is to count as a threat. Of course, we would then need to address the very difficult problem of where and how to fix the baseline.

Note, however, that we may circumvent this entire set of problems, since what concerns us here is *coercion*, not *threats*. The question whether or not the (dis)incentives under discussion are best conceptualized as threats—and thus the question of how to distinguish threats from cost-adding proposals that are not threats—would require our attention only if we accepted the idea that all threats were coercive. But this is false: On any promising account of coercion (or so I believe), whether a proposal is coercive or not depends on whether the (dis)incentives attached to it are of such magnitude as to apply on its recipient an amount of pressure sufficient to eliminate as a genuine option the activity that the proposal aims to discourage. A proposal fails to be coercive, then, if the attached (dis)incentives are too insignificant to apply the requisite amount of pressure. It is a question, of course, what options we may count as 'genuine'—i.e., as genuinely refusible, as something that someone really could choose to accept or decline, etc. In particular, it is a question whether we should think about the refusibility of options in purely descriptive or at least partly in normative terms. But whatever view we take, it does not affect my point. If what I've said so far is correct, then the question we should be asking is *not* whether perfectionist (dis)incentives count as threats, but simply whether they need be coercive—to which the plain answer is no. Any plausible view of the degree of pressure that a proposal need apply in order to count as coercive would not designate our normal subsidization and taxation practices—say, a \$1 tax on an otherwise \$10 activity—as applying coercive pressure on those they address.

¹⁵ Waldron 1989, 1142.

¹⁶ Waldron's remarks, I take it, are intended to apply to activity-specific taxation and to subsidies alike: that is, they seem to count as coercive not just threats but (at least some) offers as well.

§5 Manipulation and Inappropriate Motivation

5.1 Manipulation

There remains the possibility that trying to influence people's choices by altering the incentive structure of the options available to them might be objectionable for reasons other than coercion yet normatively very similar to it. One way of formulating the revised complaint is this: The decision to pursue an activity or to avoid it ought to be based on a certain type of consideration—namely, on considerations stemming from the activity's intrinsic value, from its intrinsic merits and demerits. 'Artificially' to add other considerations to the deliberative balance is *manipulative*, the objection goes, even if not coercive. And manipulation, while not identical to coercion, is objectionable for the same reason coercion is: it diminishes or violates the autonomy of its victims.

How exactly might economic (dis)incentives be construed as manipulative? Here, for example, are a few ways in which Waldron spells this out. Activity-specific taxation, he writes, "is necessarily manipulative, for it influences a person's decision by distorting the individual's understanding of the merits of the choice." The same is said to apply to subsidies, which "give those who benefit from [them] a misleading and distorted picture of the real costs and benefits of engaging in the subsidized activity." These policies thus "interfere with the way people form their beliefs about value"; they undermine one's ability to "respond to value by choosing for reasons the person apprehends." Finally, by "messing with the options that one faces," manipulation undermines one's ability to choose among options "for the right reasons."¹⁷

Manipulative intervention motivates people for the wrong reasons, then, and the above remarks focus on one technique in which this can be done, namely *deception*: manipulative acts can exert influence by concealing the appropriate reasons (reasons stemming from an option's real value), or by presenting inappropriate reasoning as appropriate.

But need perfectionist policies, if they are to succeed, distort people's perceptions of the genuine merits or of the "real costs and benefits" of an activity?¹⁸ The answer, I think, is straightforward: perfectionist policies need not have this epistemic effect. Although they could (if so designed) distort the information people have about the desirable and undesirable features of various options, they need not do so in order successfully to encourage or discourage the options. (Certainly, the epistemic concern can serve as no objection to measures—e.g., state-sponsored media campaigns—the entire point of which is to inform people of the value or disvalue of an option.)

Not all manipulation relies on deception, however: one could be manipulated into choosing for the wrong reasons while not for a moment losing sight of the right ones. Manipulation, as we normally understand the term, can also consist in the exploitation of a person's weaknesses, with no epistemic component involved. To respond fully, then, to the charge that economic (dis)incentives are manipulative, we will have to consider whether (dis)incentives might manipulate in this second way.¹⁹ To make the idea clearer, we will need to

¹⁷ Waldron 1989, 1145, 1148-49.

¹⁸ Note that there is a distinction to be made here (which Waldron does not acknowledge) between the *merits and demerits* of an activity and the *costs and benefits* associated with it. I will return to this distinction shortly.

¹⁹ This understanding of manipulative acts as either deceptive or exploitative of weaknesses is roughly in line with Joel Rudinow's (in my view largely successful) account of manipulation (Rudinow 1978, 338-34). In this paper I do not attempt anything approaching an originally developed or comprehensive account of manipulation. Nor do I assess in any considerable detail Rudinow's account, which, while capturing in my view the main features that

spell out in a bit more detail what exploiting someone's weakness consists in (and how influencing someone in this manner differs from manipulation by deception, as well as from coercion and from rational persuasion). I see two ways of doing so, one focused on some *peculiar facts* about the manipulated party, the other on the *degree of pressure* that the manipulative action exerts. That is, (1) we can sometimes say that A is exploiting some weakness of B when A offers B a (dis)incentive which, though not normally irresistible (i.e., not irresistible to people in general), *is* irresistible to B due to some distinctive or idiosyncratic fact about, e.g., B's personality or psychological makeup. Or (2) perhaps we can sometimes say that A is trying to play on some weakness of B when A offers B a (dis)incentive which B is expected to find very hard to resist, although the degree of irresistibility is not quite of such magnitude as to completely undermine B's ability to make a genuine choice.²⁰ [EXAMPLE]

This account is extremely sketchy, of course, and would need to be supplemented by further conditions in order to pick out manipulative actions only (and exclude, e.g., acts of coercion). And of course, we should expect the boundaries between coercion and weakness-exploiting manipulation to be somewhat vague. But it seems to me that with some fine-tuning, (1) *idiosyncrasy* and (2) *degree of irresistibility* are the two features we should focus on in order to carve out conceptual space for a class of manipulative actions that are neither deceptive (as other manipulative actions are) nor simply coercive.

It is hard to see, however, why either of these two features must be present in such perfectionist policies as subsidization or activity-specific taxation. The policies need not be designed to exploit any peculiar weaknesses of the people whose behavior they hope to influence. And the (dis)incentives involved need not be especially hard to resist. Although the point of such policies *is* to change the balance of costs and benefits attached to various options, such changes need not be dramatic: sufficiently mild (dis)incentives can avoid applying great (and perhaps morally excessive) pressure on those who strongly resist the encouragement or discouragement they offer, yet have the desired influence on many other people who are less deeply resistant.

5.2 *Inappropriate Motivation*

The charge that perfectionist (dis)incentives are necessarily manipulative may be unfounded, but it rests on an even more basic worry which, in its most general form, has not been addressed yet—namely, the worry that perfectionist (dis)incentives might motivate people to act *for the wrong reasons*. That is, even if a (dis)incentive is not manipulative in any ordinary sense of the term—even if it is neither deceptive nor exploitative in the above-discussed sense—it might still motivate inappropriately, and *that*, goes the objection, is still sufficiently troubling.

The worry about inappropriate motivation—as formulated, for example, in some of the above excerpts from Waldron—depends on the persuasiveness of the distinction between an option's intrinsic merits (which provide the 'right' reasons for choosing it), on the one hand, and factors introduced to it 'artificially' or 'from the outside', on the other. Now, let us assume for a moment that the distinction is unproblematic. (I will return to this assumption in a moment.) Even so, it is unclear why considerations shaped by state action should fall on one side of the line, and considerations shaped by other social forces on the other. Why should incentives

distinguish manipulation both from rational persuasion and from coercion, is incomplete in certain ways, particularly with regard to the notion of 'playing on [or exploiting] someone's weakness'.

²⁰ For a discussion of the first, 'personality'-based condition, see Rudinow 1978, 342.

designed and administered by the state be considered any more extraneous to the value of an activity than incentives created and sustained by other, non-state agents? Consider a composer who needs to decide whether to devote her time to composing concert music or to recording advertisement jingles. One class of considerations in favor of the former option is that it is more aesthetically valuable, more personally fulfilling, more conducive to the full realization of one's talents, etc.; a different kind of consideration might be that the state offers generous grants to composers on the basis of artistic merit. Meanwhile, a consideration in favor of the latter option might be that advertising companies pay extremely well. Surely, if there is a line to be drawn here between two different groups of considerations, it should separate the first consideration from the latter two, not the second from the third.

Recall the condition that anti-perfectionist arguments must satisfy: they must explain why it is the *state* that ought to banish perfectionist considerations from the justification of its actions. But this is precisely what the argument from inappropriate motivation has failed to do so far: any restrictions it might justify would apply to state and to non-state action alike.²¹

I think we can do better, however, than this sort of response (which concedes, after all, that (dis)incentives 'motivate for the wrong reasons', merely adding that it is no more objectionable when the state uses them than when non-state agents do). To see how, let us return to the distinction between the 'intrinsic' and the 'extraneous' advantages and disadvantages of an option.

One possible response to the 'inappropriate motivation' objection is, of course, to deny that the distinction can, or can typically, be made—to insist (as, e.g., Raz does in response to Waldron's articulation of the objection) that the notion of "intrinsic merit ... independent of social conditions in most cases is hard to sustain"²²—and thereby to undermine the basis for a demarcation between 'appropriate' and 'inappropriate' motivation.

This seems too strong, however. Returning to our composer example, I do not see a difficulty in distinguishing, as I have, between two classes of considerations on which the composer's choice could be based: the intrinsic merits or demerits of each of the two options, and (in this example) the financial considerations attached—but extraneous—to each option.²³

But we need not reject the distinction in order to respond to the inappropriate motivation objection. The assumption that we do rests, I think, on a conflation of *two* distinctions that the objection in fact involves: (1) a distinction between the intrinsic (de)merits of an option and the associated-yet-extraneous costs and benefits; and, (2) *within* the extraneous costs and benefits, a distinction between 'artificial' and 'authentic' cost/benefit balances. We should not conflate, however (as Raz's response seems to do), the claim that the notion of an option's intrinsic merits is unsustainable with the very different claim that there might be no way of demarcating an option's 'authentic' or 'natural' cost/benefit balance. And if that is correct, then it remains possible—in fact, I think it is the case—that (in a great many instances at least) we can distinguish between an option's intrinsic (de)merits and the extraneous costs and benefits attached to it, even though within the latter category there is no way of distinguishing between 'artificial' cost/benefit balances and ones that are 'natural', 'authentic', or uninterfered-with.

²¹ Cf. Raz 1989, 1234; Wall 1998, 200.

²² Raz 1989, 1234.

²³ Perhaps the distinction cannot be made in all cases; and perhaps it is more difficult to make when what we have on the supposedly 'extraneous' side are not financial considerations but, say, considerations of social esteem. But that is less significant for our purposes, for what concerns us here are precisely financial (dis)incentives.

It is very rarely, if ever, of course, that we can make choices without taking extraneous (at the very least, financial) considerations into account. Our choices are normally motivated by a mixture of such considerations, on the one hand, and the perceived merits of the options before us, on the other. The more, however, we are guided by an appreciation of the intrinsic (de)merits of options rather than by the extraneous (dis)incentives attached to them, the better. And while I do not think we can distinguish between ‘authentic’ and ‘artificial’ cost/benefit balances (surely we should not associate the former with market- and the latter with government-regulated ones!), we *can* distinguish between cost/benefit balances that make it easier for us to act for the right reasons—to base our decisions on the intrinsic merits of the options before us—and ones that diminish our ability to do so.

Where does all this leave us with respect to the perfectionist policies we’ve been discussing? In some cases, the use of economic incentives clearly *lacks* the effect of adding extraneous considerations to the overall deliberative balance (or of making such considerations weightier): on the contrary, they can enable the recipients to be *more* rather than less responsive to the intrinsic merits of the options before them. This is the case when subsidizing an activity makes financial considerations loom less large in people’s deliberations and thus frees them to focus on and respond to the activity’s value. (This effect is especially clear when those affected by a subsidy are eager to engage in the subsidized activity anyway, but would be deterred by its cost if not for the subsidization.) In other cases, those who might be affected by the subsidy (and this applies equally to taxation) might initially be indifferent to the encouraged option, and, insofar as they choose it, might indeed choose it initially for the ‘wrong’ reasons—because it is inexpensive enough to be chosen on a whim, and so forth. In such cases we may hope, however, that exposure to the encouraged option will convince the initially indifferent of its merits, and that later engagement will therefore take place for the right reasons.²⁴

But what about the problematic cases, those in which people want to pursue a taxed activity (or, perhaps less problematically, do not want to pursue a subsidized one?) Well, to the extent that we do not want people to be motivated by the wrong considerations, the answer is that government (dis)incentives should be set at such a level as to affect the first two groups of people—the eager and the indifferent—and not the latter—that is, the resistant.²⁵ That is the concession that perfectionist (dis)incentives would have to make if the worry about inappropriate motivation proved to be sufficiently troubling.

§6 Conclusion

I have argued that instruments of perfectionist policy such as economic (dis)incentives can leave people perfectly free to decide which conceptions of the good to endorse and in which activities to engage, and thus, in at least this crucial respect, need not diminish their autonomy. Perfectionists can—and, I think, should—also argue, however, that some measure of freedom or autonomy may sometimes be sacrificed for the sake of other goods, especially when the policies in question involve forms of pressure much milder than coercion. But even short of such arguments, I hope to have shown that appeals to the state’s coercive character cannot ground a thoroughgoing case against political perfectionism. There may, of course, be other general

²⁴ Cf. Raz 1989, ???; Kymlicka 1989, ???; Sher 1997, 64; Wall 1998, ???.

²⁵ Alternatively, we could hope that those who at first strongly resist an option might later, as a result of some incentive, come to see its merits. I think it is obvious, though, that such hope will not be realistic in all cases.

features of the state besides its coercive character—its symbolic significance, perhaps—in virtue of which special blanket restrictions might apply to the justification of its actions. But arguments that invoke such further features should not be conflated with arguments from coercion, and would pose, I believe, less of a threat to political perfectionism.

REFERENCES

- Ackerman, B. (1980). *Social Justice in the Liberal State*. New Haven: Yale University Press.
- Gaus, G. (2003). "Liberal Neutrality: A Compelling and Radical Principle," in Steven Wall and George Klosko (eds.), *Perfectionism and Neutrality: Essays in Liberal Theory*. Lanham: Rowman & Littlefield.
- Hurka, T. (1993). *Perfectionism*. Oxford: Oxford University Press.
- Kymlicka, W. (1989.) "Liberal Individualism and Liberal Neutrality." *Ethics* 99:4 (1989), 883-905.
- Kymlicka, W. (1990). *Contemporary Political Philosophy*. Oxford: Oxford University Press.
- Larmore, C. (1987). *Patterns of Moral Complexity*. Cambridge: Cambridge University Press.
- Nagel, T. (1987). "Moral Conflict and Political Legitimacy." *Philosophy and Public Affairs* 16:3, 215-240.
- Rawls, J. (1971). *A Theory of Justice*. Cambridge: Belknap, 1971.
- Rawls, J. (1993). *Political Liberalism*. New York: Columbia University Press, 1993.
- Raz, J. (1986). *The Morality of Freedom*. Oxford: Oxford University Press.
- Raz, J. (1989). "Facing Up." *Southern California Law Review* 62, 1153-1235.
- Rudinow, J. (1978). "Manipulation." *Ethics* 88:4, 338-347.
- Sher, G. (1997). *Beyond Neutrality*. Cambridge: Cambridge University Press.
- Waldron, J. (1989). "Autonomy and Perfectionism in Raz's *Morality of Freedom*." *Southern California Law Review* 62, 1097-1152.
- Waldron, J. (1993). "Legislation and Moral Neutrality," in *Liberal Rights*. Cambridge: Cambridge University Press.
- Wall, S. (1998). *Liberalism, Perfectionism and Restraint*. Cambridge: Cambridge University Press.