

Anti-Perfectionisms and Autonomy

Ben Colburn

Published in *Analysis* 70 (2010): 247-256.

Different types of liberal theory can be distinguished according to their commitments in respect of the following two claims:

The Autonomy Claim: The state ought to promote autonomy.

Anti-perfectionism: The state ought not in its action to promote any value.

Autonomy, as it will be understood here, is a value which consists in an agent deciding for herself what is a valuable life, and living her life in accordance with that decision.¹ Anti-perfectionism, in this context, is the thesis that the state ought not in its actions intentionally to promote any value or putative value.

Some liberals – call them *perfectionistic liberals* – are committed to the Autonomy Claim, and to the denial of Anti-perfectionism.² Others – call them *political liberals* – have the converse commitments: they endorse Anti-perfectionism, and deny that the state ought to promote autonomy.³ Yet other liberals avoid both commitments, rejecting the state's promotion of autonomy but endorsing the promotion of some other value.⁴

In this paper, I provide support for a liberal political philosophy that endorses both of the claims given above: that is, one that is fully committed to the state promotion of autonomy, and which also counts Anti-perfectionism amongst its other commitments.⁵ I do so by defending it against the serious charge that it is *prima facie* self-contradictory. After all, Anti-perfectionism appears to demand that the state refrain from promoting any value – it looks as though that must preclude the promotion of autonomy, if the latter is conceived of as a value. I argue that this self-contradiction is a mirage, whose plausibility depends on an equivocation in the statement given above of Anti-perfectionism. As it stands, the statement fails to distinguish between two types of values, and hence between two different types of Anti-perfectionism that one might endorse. When this is clarified, we can see that on the best understanding of Anti-perfectionism, it is consistent with the Autonomy Claim.

¹ Understanding autonomy in this way puts me in the company of many people who have written about autonomy as a foundation for liberalism: most notably Joseph Raz, who takes autonomy to be an 'ideal of self-creation', and of an agent being part author of her own life (Raz 1986: p.369). I set out in more detail what this conception of autonomy amounts to in Colburn 2010: chapter 2.

² e.g. Raz, and also Wall (1998) and Hurka (1993).

³ Rawls' later work implies anti-perfectionism, and he often said that this commitment implies the rejection of state promotion of autonomy. See, for example, Rawls (1985) and (1993). Rawls' intellectual successors within political liberalism (e.g. Macedo 1990) take the same view.

⁴ e.g. William Galston, who defends diversity-minded liberalism (Galston 1995).

⁵ This possibility has been little discussed, possible because of the whiff of contradiction that the present paper aims to dispel. See, however, Oshana 2006: chapter 5, and Clayton 2006 pp. 11 – 19.

1 The idea of a variable

It is consistent to affirm both Anti-perfectionism and the Autonomy Claim because Anti-perfectionism is ambiguous between two readings, only one of which is incompatible with the state's promotion of autonomy. This section and the next one demonstrate this by distinguishing between two types of value, called 'first-' and 'second-order' values. So long as we understand Anti-perfectionism as concerning the state promotion of *first-order* values, and we take autonomy to be a *second-order* value, then the two claims are compatible.

To draw the distinction between first- and second-order values, I first need to introduce the concept of a *variable in the specification of a value*. This is easiest done in the context of the more familiar distinction between content-specific and content-neutral values, and so I start with a discussion of that, suggesting a new way in which the distinction can be characterised. Doing so will have an ancillary expository benefit too: it will show that, while the crucial first-order/second-order distinction is *related* to the content-specific/content-neutral distinction, the two are not the same.

Suppose that I ask two people what is valuable in life, and that they answer as follows:

- Anne says: 'All that is valuable in life is to be able to play Bach's Cello Suites flawlessly.'
- Barbara says: 'What is valuable in life is satisfaction of desire.'

Suppose also that Anne's statement is to be taken at face value — that is, she believes playing Bach's Cello Suites flawlessly to be intrinsically and non-derivatively valuable (as opposed to thinking just that it is a good means to promoting some other value like happiness, for example). Anne and Barbara's claims illustrate a distinction between two types of non-derivative value: Anne posits a *content-specific* value, and Barbara a *content-neutral* value.

A content-specific value is one that fully specifies some states of affairs as what are valuable. Anne posits a value which is instantiated just when people are playing Bach's Cello Suites. Nothing further need be said to identify states of affairs which I should promote in order to ensure that, by Anne's lights, my life goes well: I just need practical information about how to promote those states of affairs.

A content-neutral value, on the other hand, does not fully specify any particular states of affairs. Barbara, recall, says that desire-satisfaction is valuable. To pursue *that* value, or to evaluate someone else's life with respect to it, I need to know about my desires and what will satisfy them. Without this information, the specification of a content-neutral value is incomplete. This incompleteness is the characteristic feature of content-neutral, as opposed to content-specific values. Precisely because the value is sensitive to individual attitudes, we cannot specify it any more determinately (in terms of states of affairs) than by saying that desire-satisfaction is valuable.

Moreover, this incompleteness is ineliminable: it cannot be translated away by substitution. To show this, consider what happens if one tries to eliminate the incompleteness in the specification of a content-neutral value by replacing the problematic term with something that fully specifies what is valuable. Suppose that I desire tea. Could we substitute 'tea' for 'desire-satisfaction' in Barbara's statement, and thereby give a complete specification of Barbara's value by saying 'tea is valuable'?

The answer is: plainly not, for two reasons. First, the translated sentence does not mean the same as Barbara's original claim. She said that it is true of everyone that desire-

satisfaction is valuable, but she is not committed to the same claim about tea. At best we have added an implicit clause relativising the claim so that it just applies to our subject ('What is valuable *for me* is . . .'), but in so doing we have replaced a universal claim with a non-universal one, and a translation that changes the truth value of a claim is hardly an acceptable one. It should be clear that any similar substitution would result in a similar mistranslation. Secondly, the substitution involves mislocating the value. According to Barbara, desire-satisfaction *really is* valuable, and tea only derivatively valuable insofar as it satisfies our subject's desire. Putting it the other way round places priority on the wrong thing, for Barbara's reference to what is desired is *de dicto*, rather than *de re*. What is valuable is the satisfaction of desires, whatever they happen to be: it just happens, for me, now, that tea will serve.

So, to recap, content-neutral values differ from content-specific ones in that their specifications are ineliminably incomplete. To aid clarity, let us say that such a specification contains a *variable*. In different contexts, different things can instantiate the variable. For a desire-satisfaction theorist it will be that which satisfies a particular agent's desires. For a hedonist, it will be that which brings a particular individual pleasure.⁶ However, the variable cannot be eliminated (that is, be permanently replaced with something specific) without a change of meaning or truth value. Content-specific values are those values which contain no variables; content-neutral values are those which do.

2 First- and second-order values

The notion of a variable allows us to draw the distinction needed to prove the consistency of Anti-perfectionism and the Autonomy Claim, namely the distinction between *first-* and *second-order* values. These differ in respect of the types of variables they can contain.

Compare Barbara's advice — that desire-satisfaction is valuable — with advice from Carol, who says that it is valuable to follow your parents' values. Both values are content-neutral: just as Barbara's advice makes *de dicto* reference to those things that satisfy desires, Carol's advice makes *de dicto* reference to one's parents' judgments of value. However, they make different types of value claim. Barbara's advice is incomplete because we need to know what our desires are and what satisfies them: that is, states of affairs such as 'X's drinking tea' or 'X's playing the Cello Suites'. Carol's statement, by contrast, tracks other specifications of value, namely those which one's parents would give: 'Desire-satisfaction is valuable', perhaps, or 'Pleasure is valuable'. That second claim about values will determine which states of affairs someone should pursue, if they are trying to seek that value themselves, or look for, if they are trying to assess how well other people's lives exemplify it. So, here is a second distinction. Some values are specified in such a way as to refer *de dicto* to other specifications of values. Others do not: either they attach value to states of affairs *de re* (and are content-specific), or they refer to states of affairs *de dicto* (and are content-neutral).

These values differ because their variables have different ranges. The variable in Barbara's specification tracks states of affairs that satisfy desires; while that in Carol's ranges over other specifications of values. Call these two types of variables *first-* and *second-order*

⁶ Assuming, that is, that we can understand hedonism as content-neutral.

respectively.⁷ Then, define second- and first-order *values* as follows. Second-order values are those which contain a second-order variable (that is, refer *de dicto* to other specifications of value). First-order values are those which don't: either because they attach value to states of affairs *de re* (and so contain no variables at all), or because they refer to states of affairs *de dicto* (and contain only first-order variables).⁸

3 Establishing Consistency

The distinction between first- and second-order values and the observation that autonomy is second-order give me the wherewithal to prove the consistency of the two claims I set out at the start of this paper.

First, in light of the distinction, we can now see that autonomy is a *second-order* value, as least when autonomy is conceived of in the way that I suggested above.

Corresponding to the distinction between first- and second-order values one can find in the literature two different notions of autonomy. Sometimes, writers take autonomy to be an ideal of a particular way of life: most frequently one of relentless Socratic questioning, of perpetual reassessing of motives and desires and of careful, rational action.⁹ At other times, autonomy is taken to be characterised precisely by not specifying any particular way of life that must be led in order for an individual to be autonomous. So, for example, Stephen Wall says that autonomy involves 'making something out of their lives according to their own understanding of what is valuable and worth doing' (1998: p. 128). A similar idea seems to lie behind Thomas Hurka's formation that to be autonomous 'is to direct oneself where different directions are possible' (1987: 361), and Joseph Raz's ideal of self-creation and self-authorship (1986: p. 370). All three of these writers take it to be crucial to autonomy that an individual not only lives in a particular way but also that they themselves endorse the standards they use to assess the various options that are open to them.¹⁰

These two ways of talking about autonomy differ in referring to two different types of values. There may be a single concept at work, but it is liable to two different interpretations. The first refers to a first-order (and, indeed, content-specific) value: there is a particular set of things that one must do in order to be autonomous. One must go through a particular set of mental processes, perhaps, such as those demanded by the person who advocates a life of Socratic self-questioning.

⁷ For the sake of simplicity, the examples used here all involve reference to someone else's judgments of value. It is easier to highlight the relevant features of second-order variables by considering these sorts of cases, but the reference to third parties is not essential.

⁸ So, all second-order values are content-neutral, but some content-neutral values — namely those which contain only first-order variables — are first-order.

⁹ e.g. Cooper 2003, Ladenson 1975, and Oshana 2003.

¹⁰ This might be thought to ignore Raz's insistence on the importance of choosing between *valuable* objects, where what counts as valuable is antecedently given (1986: pp.289–320). *That* claim, though, belongs more properly to Raz's views on well-being, which ought to be distinguished from his theory of autonomy: deciding for oneself what is valuable is at the core of Raz's view of autonomy, even if it is only a necessary and not a sufficient condition for one's autonomy contributing to one's well-being (1986: p. 381).

The second interpretation, by contrast, refers to a second-order value. Such is the conception of autonomy on which this paper is based, according to which autonomy consists in an agent deciding for herself what is a valuable life and living her life in accordance with that decision. In this specification some conditions upon autonomy are given, but there is an ineliminable variable which stands for an individual living the sort of life that he deems valuable: one cannot specify in any more definite way how people must live in order for their lives to be autonomous. Moreover, the variable is a second-order one. The specification refers to other judgments of what is valuable; namely, to those which an individual decides on for herself. Moreover, the reference is *de dicto*: taking the reference as *de re* (and thereby allowing that reference to be replaced by a statement of what the individual does, in fact, consider valuable) would change the meaning of the claim and locate what is valuable in the wrong place. Therefore, autonomy on this understanding is second-order, since its specification must be able to contain a second-order variable.

The distinction between first- and second-order values also reveals an ambiguity in the way that I originally formulated Anti-perfectionism. When one affirms Anti-perfectionism, one might be making a claim about first-order values, or second-order values, or both, or neither. Let us then remove the equivocation by saying that we shall understand it as *first-order Anti-perfectionism*: that is, as the claim that the state ought not intentionally to promote *first-order* values. Since autonomy is a second-order value, our new construal of Anti-perfectionism is consistent with the Autonomy Claim, and the thesis of this paper established – so long as so it is fair to construe Anti-perfectionism in this way.

4 Varieties of Anti-perfectionism

To recap: Anti-perfectionism as I defined it at the start of this paper is ambiguous. My proposal is that we disambiguate it by specifying that it ranges over first-order values. In this section I address two worries that one might have with this move.

To start with, one might worry that disambiguating Anti-perfectionism in the direction of only first-order values is *ad hoc* – purely motivated by the desire to invent a position consistent with the Autonomy Claim (which concerns the second-order value of autonomy).

The charge of being *ad hoc* is mistaken, because those philosophers who endorse Anti-Perfectionism almost invariably have the state promotion of first-order values as their target. To take a single (but very influential) example, in *Political Liberalism* Rawls gives three examples of comprehensive doctrines on values which the state ought not, in his view, to promote: orthodox medieval Catholicism, utilitarianism and the ‘liberalisms of Kant or Mill’, based on ideals of individuality or reason (1993: p. 37). That which Rawls identifies as impermissible in each case is commitment to values which are first-order, as the term was defined above.

The second worry is more serious. It may be that the preceding observations suffice to show that understanding Anti-perfectionism as *first-order* Anti-perfectionism is not contrary to the letter of a theory like Rawls’. But it might be objected that my proposal nevertheless fails to capture the spirit of Anti-perfectionism. There is a whole host of putative second-order values, the state promotion of which would be just as offensive to anti-perfectionist intuitions. In some cases, this will be because – in addition to the *de dicto* reference to another specification of value – a second-order value can contain as much unwelcome specificity about someone’s way of life as a first-order one. (Consider, for example, someone who said ‘It is valuable to live the life of a slave except for one day a year

when one should do as one deems valuable'.) Or, it may just be that the presence of a second-order value doesn't prevent the way of life involved from potentially being a repressive and unattractive one. Carol's insistence on following one's parents' values is a prime example.

So, one might argue, while it is logically possible to be a first-order but not a second-order anti-perfectionist, anyone who is motivated to adopt the former position should also adopt the latter. And if that is so, the formal consistency of the Autonomy Claim with first-order Anti-perfectionism does not suffice to show that there is space in the territory for a political theory committed to both.¹¹

The objection fails, for two reasons. One has to do with the underlying motivation for endorsing any form of Anti-perfectionism in the first place. I have argued elsewhere (Colburn 2010: § 3.3) that the only credible motivation for first-order Anti-perfectionism is either an explicit endorsement of or tacit reliance on the Autonomy Claim. Hence, someone who thinks that first-order Anti-Perfectionism is true cannot reasonably think the same of its second-order cousin, since the latter position is inconsistent with the only credible motivation for the former. So, the unpersuaded anti-perfectionist has to choose between first- and second-order Anti-perfectionism – and since (as I noted above) the paradigm examples of impermissible state action given by anti-perfectionists overwhelmingly involve first-order values, it seems reasonable to predict that they will settle on first-order Anti-perfectionism out of the two.

The second reason that the objection fails is this: it relies on the assumption that restricting oneself to first-order Anti-perfectionism is unstable because it must offend the intuitions that make any form of Anti-perfectionism attractive in the first place. In particular, I suppose the worry is that restricting oneself to first-order Anti-perfectionism means having to refrain from condemning state actions which seek to promote second-order values.

Plainly, though, that is not the case. For one thing, if one endorses the Autonomy Claim in addition to first-order Anti-perfectionism, then that by itself gives ample reason to think that many such actions are impermissible. If the state should promote individual autonomy, then it certainly shouldn't be in the business of trying to get people to do only what their parents consider valuable, for example. Indeed, this reasoning might well rule out *all* instances of second-order perfectionism save for the Autonomy Claim itself. So, in practice the autonomy-minded anti-perfectionist could be – *should* be – an anti-perfectionist about almost everything anyway, second-order values included.

Even if one rejects the Autonomy Claim,¹² restricting oneself to full-blown Anti-perfectionism about only first-order values doesn't mean thinking that any state action seeking to promote a second-order value is acceptable. It is quite consistent with denying the universal claim of second-order Anti-perfectionism to think, of any instance of second-order perfectionism one might care to mention, that there are good reasons to oppose it. Therefore the objection fails to establish that the anti-perfectionist's intuitions must be offended by the restriction to first-order Anti-perfectionism – in which case, my proposal that we do so is undefeated.¹³

¹¹ My thanks to an anonymous referee for this objection.

¹² As noted above, I think rejecting it is inconsistent with first-order Anti-perfectionism, but I shan't assume so here.

¹³ It might be argued that this still won't do justice to the anti-perfectionist's intuitions, because those intuitions are not merely that such-and-such an instance of second-order

Conclusion

Having drawn the distinction between first- and second-order values, I have been able to establish two things. First, autonomy should be understood as a second-order value. Secondly, on a reasonable understanding of the position, Anti-perfectionism should be understood as *first-order* Anti-perfectionism. So, Anti-perfectionism and the Autonomy Claim do not contradict each other. Hence, a liberal political theory which is committed to both is consistent, and the appearance of paradox which I observed at the start of this paper is dissolved.¹⁴

References

- M. Clayton. 2006. *Justice and Legitimacy in Upbringing*. New York: Oxford University Press.
- B. Colburn. 2010. *Autonomy and Liberalism*. New York: Routledge.
- J.M. Cooper. 2003. Stoic Autonomy. *Social Philosophy & Policy* 20 (2): 1 – 29.
- W. Galston. 1995. Two Concepts of Liberalism. *Ethics* 105: 516 – 534.
- T. Hurka. 1987. Why value autonomy? *Social Theory and Practice* 13: 361 – 380.
- T. Hurka. 1993. *Perfectionism*. New York: Oxford University Press.
- R.F. Ladenson. 1975. A theory of personal autonomy. *Ethics* 86: 30 – 48.
- M. Oshana. 2003. How much should we value autonomy? *Social Philosophy and Policy* 20 (2): 99 – 126.
- M. Oshana. 2006. *Personal Autonomy in Society*. Aldershot: Ashgate.
- J. Rawls. 1985. Justice as Fairness: Political not Metaphysical. *Philosophy and Public Affairs* 14: 233 – 251.
- J. Rawls. 1993. *Political Liberalism*. New York: Columbia University Press.
- J. Raz. 1986. *The Morality of Freedom*. Oxford: Clarendon Press.
- S. Wall. 1998. *Liberalism, Perfectionism and Restraint*. New York: Cambridge University Press.

perfectionism is wrong, but that it is wrong *because it is an instance of the general principle of second-order Anti-perfectionism*. Now, I strongly doubt that intuitions have such sophisticated content, but the point need not be argued here. I shall just note that the objection has shifted from saying that second-order Anti-perfectionism must be correct because the intuitions that support first-order Anti-perfectionism support it too, to saying that second-order Anti-perfectionism must be correct because we have an intuition that second-order Anti-perfectionism is correct. And that is plainly question-begging.

¹⁴ My thanks to Hallvard Lillehammer, Serena Olsaretti, Jonathan Wolff, Matthew Clayton, Daniel Elstein, Brian King, Neil Sinclair and David Liggins for helpful comments and criticisms of this material; and also to audiences in Bristol, Cambridge, Montreal, Pavia and Warwick, where earlier versions were aired. I gratefully acknowledge the support of an AHRC doctoral grant during the writing of this paper.