

- Unger, P. 1975. *Ignorance: A Case for Scepticism*. Oxford: Clarendon Press.
- Vogel, J. 1990. Cartesian skepticism and inference to the best explanation. *Journal of Philosophy* 87: 658–66.
- Williamson, T. 2000. *Knowledge and its Limits*. Oxford: Oxford University Press.
- Wright, C. 1991. Scepticism and dreaming: imploding the demon. *Mind* 100: 87–116.

## Take the sugar

CASPAR HARE

Sometimes I lack all-things-considered preferences between items. Sometimes this attitude is *insensitive to mild sweetening*. There are items A, A+, B, B+, such that, all things considered, I have no preference between A and B, I have a preference for A+ over A, I have a preference for B+ over B, and yet I have no preference between A and B+, or between A+ and B.

The attitude may be turbulent, naturally described as one of ‘tortured ambivalence’:

### *The fire*

Firefighters are retrieving possessions from my burning house. Should I direct them towards the Fabergé egg in my drawing room or the wedding album in my bedroom? The Fabergé egg was commissioned by Czar Alexander III of Russia, as an Easter surprise for the Empress Consort. It has survived revolution, war and upheaval on a grand scale, and is now regarded as the finest relic of the gaudy, opulent Romanov dynasty. The wedding album, on the other hand, reminds me of happy times when my wife and I were young and careless. As I think, in turn, of losing the one or the other, my emotions and inclinations vacillate wildly, never settling down to the point where it would be fair to describe me as having an all-things considered preference between:

A: The firefighters saving the Fabergé egg.

and

B: The firefighters saving the wedding album.

Force me to choose and I will choose. But my choice will have an arbitrary flavour. And learning that there is a \$100 bill lying beside the egg or the album will not rid it of this flavour. When I compare B to:

A+: The firefighters saving the Fabergé egg, plus \$100.

and A to:

B+: The firefighters saving the wedding album, plus \$100.

I remain just as ambivalent as before. I have no all-things-considered preference between A+ and B, B+ and A, though I do prefer A+ to A, B+ to B.

Or the attitude may be calmer, more naturally described as one of ‘indifference’:

*The dinner*

It is dinner-time. Should we go to the Indian restaurant or the Chinese restaurant? We have visited both many times. We know their pluses and minuses. The Indian restaurant is less far to walk. It serves up a sublime mango lassi. The Chinese restaurant is cheaper. Its raucous atmosphere is more child-friendly. All in all it is a wash for me. I have no all-things-considered preference between:

A: Our going to the Indian restaurant.

and

B: Our going to the Chinese restaurant.

And learning that it is dollar-off day at either restaurant will not give me an all-things-considered preference. When I compare B to:

A+: Our going to the Indian restaurant and saving \$1.

and A to:

B+: Our going to the Chinese restaurant and saving \$1.

it remains a wash for me. I have no all-things-considered preference between A+ and B, B+ and A, though I do prefer A+ to A, B+ to B.

This isn’t just me. I take it that we all have patterns of preference like this, all the time. Indeed, a major project in recent normative theory has been to develop a view according to which patterns of preference like this are appropriate responses to real evaluative relations between items.<sup>1</sup>

1 Briefly: On one view, in these cases there’s a relevant way of being better such that A+ is better than A, and it is not the case that A+ is better than B or vice-versa, and it is not the case that A is better than B or vice-versa. A and A+ stand in a strange evaluative relation to B. It would be misleading to call the relation ‘being equally good’, because that suggests transitivity, and the relation is not transitive. James Griffin called it ‘rough equality’ – see Griffin 1986: 80–1, 96–8 and also Parfit 1984: 431–2. Ruth Chang calls it ‘parity’ – see Chang 2002 and 2005. On another view, in these cases there’s a relevant way of being better such that A+ is better than A, and it is indeterminate whether A+ is better than B or vice-versa, and it is indeterminate whether A is better than B or vice-versa. John Broome has carefully developed and defended this view in Broome 1997 and 2000. For present purposes it will not matter which view is right.

So, how is it rational for me to proceed when I have such preferences, and various options are open to me and I am unsure about what will happen if I pursue them? This looks like the sort of question that theories of rational choice under conditions of uncertainty might help us answer. But the standard theory, the theory whose general form we have inherited from Bayes, Ramsey and von Neumann, is no help. The standard theory begins by associating a utility function with my preferences. This is a function from outcomes to numbers that (among other things) assigns a higher number to one outcome than another iff I prefer the one outcome to the other. Such a function exists only if my preferences between outcomes are *negatively transitive* (only if, for all outcomes  $x, y, z$ , if I do not prefer  $x$  to  $y$ , and I do not prefer  $y$  to  $z$ , then I do not prefer  $x$  to  $z$ ).<sup>2</sup> My preferences are negatively intransitive (I do not prefer  $A+$  to  $B$ , and I do not prefer  $B$  to  $A$ , but I do prefer  $A+$  to  $A$ ), so there is no such function.

‘So much the worse for the question’, a standard theorist might say. ‘It is not my job to guide you when you have negatively intransitive preferences. Reflect a bit, render your preferences negatively transitive, and then come back to me.’

This is unhelpful. I have no inclination to render my preferences negatively transitive. And even if I did have an inclination to do it, doing it would involve acquiring new preferences or dropping old ones. It is not so easy to acquire or drop preferences at will.

A more constructive response is to build a theory of rational decision under conditions of uncertainty to cover situations in which we have negatively intransitive preferences. In the first section of this paper I will draw attention to a problem that we face as soon as we take this project seriously. In the second I will present two theories, corresponding to different solutions to the problem.

### 1. *The opaque sweetening problem*

Suppose I lack preferences between my getting item  $A$  and my getting item  $B$ . Suppose this attitude is insensitive to mild sweetening. And suppose we play a kind of game:

#### *Two opaque boxes*

You show me items  $A$  and  $B$ , a dollar, a coin, and two opaque boxes.

You toss the coin and, governed by the toss, place item  $A$  in one box

- 2 Classic expositions of the standard theory secure the negative transitivity of preferences by way of axioms that state that weak preferences (where I weakly prefer  $a$  to  $b$  when I prefer  $a$  to  $b$  or I am indifferent between  $a$  and  $b$ ) are transitive (for all  $a, b, c$ , if I weakly prefer  $a$  to  $b$  and  $b$  to  $c$  then I weakly prefer  $a$  to  $c$ ) and complete (for all  $a, b$ , either I weakly prefer  $a$  to  $b$  or I weakly prefer  $b$  to  $a$ ). One way to accommodate the negative intransitivity of preferences is to drop the transitivity axiom, another is to drop the completeness axiom. Is it that I am indifferent between  $A$  and  $B$ ,  $A+$  and  $B$ , and my indifference is intransitive? Or is it that I have some other, *sui generis* attitude towards  $A$  and  $B$ ,  $A+$  and  $B$ ? For present purposes it will not matter how we answer these questions.

and item B in the other. I don't see which item went where. You toss the coin again and, governed by the toss, place the dollar inside the right box. I see that – which leaves me with credence 0.5 that things are like so:

Left Box	Right Box
<div style="border: 1px solid black; padding: 2px; display: inline-block;">A</div>	<div style="border: 1px solid black; padding: 2px; display: inline-block;">B + \$1</div>

and credence 0.5 that things are like so:

Left Box	Right Box
<div style="border: 1px solid black; padding: 2px; display: inline-block;">B</div>	<div style="border: 1px solid black; padding: 2px; display: inline-block;">A + \$1</div>

Then you invite me to walk away with one of the boxes.

Given what I know and prefer, what is it rationally permissible for me to do in this case? Here are two seemingly powerful arguments to the conclusion that it is rationally impermissible to take the left, unsweetened box:

**Argument 1. I have no reason to take the left, rather than the right, box**

There is a consideration of which I am aware that counts in favour of my taking the right, rather than the left box: I will get a dollar if I take the right box, no dollar if I take the left box. But there is no consideration of which I am aware that counts in favour of my taking the left, rather than the right box. So, it is rationally impermissible for me to take the left box. It is rationally impermissible to do something when I have no reason to do it, and a reason to do something else.

**Argument 2. I will improve my prospects by taking the right box**

Think of the *prospect* associated with an option as, roughly, the things I think might happen if I take it, weighted by how likely I think them to happen, if I take it. More precisely, let the prospect be the set of pairs  $\langle c, o \rangle$  such that  $o$  is an outcome that might, for all I know, come about if I take the option, and  $c$  is my credence that the outcome will come about if I take the option.<sup>3</sup> Here's a *prima facie* plausible claim about prospects and rational permissibility:

*Prospects Determine Permissibility*

Facts about what it is rationally permissible for me to do are determined by facts about the prospects associated with the options available to me.

3 If we wish to avoid committing ourselves to the denial of causal decision theory, then we should understand my 'credence that the outcome will come about if I take the option' in a particular way. It is not just my credence in the outcome, conditional on my taking the option. Rather, my credence in outcome  $o$ , relative to option  $a$ , is  $\sum_d (P(d) \cdot P(o/ad))$ , where  $d$  is a variable ranging over dependency hypotheses, ' $P(d)$ ' refers to my credence in dependency hypothesis  $d$ , and ' $P(o/ad)$ ' refers to my credence in outcome  $o$ , conditional on my taking option  $a$  and dependency hypothesis  $d$  being true. But this will not make a difference in any of the cases I discuss here.

What it is rationally permissible for me to do depends only on the things I think might happen if I take the options open to me, and how likely I think them to happen.

Now consider another game:

*One opaque box*

You show me items A and B, a dollar, a coin, and one opaque box. You toss the coin and, governed by the toss, place item A or item B in the box. I don't see which. Then you invite me to walk away with the box and the dollar, or just the box.

Obviously I have to accept the dollar in this case. But the prospects associated with the options available to me in this case are the same as the prospects associated with the options available to me in the two opaque boxes case. In this case, the prospect associated with my taking the box alone is  $\{<0.5, A>, <0.5, B>\}$  (which is to say that I think it 0.5 likely that I will end up with A, 0.5 likely that I will end up with B, if I take the box alone), and the prospect associated with my taking the box and the dollar is  $\{<0.5, A+>, <0.5, B+>\}$ . In the two opaque boxes case the prospect associated with my taking the left box is  $\{<0.5, A>, <0.5, B>\}$ , and the prospect associated with my taking the right box is  $\{<0.5, A+>, <0.5, B+>\}$ .<sup>4</sup> So, by *Prospects Determine Permissibility*, in the two opaque boxes case it is rationally impermissible to take the left box.

Is that the end of the matter – I have to take the right box? Maybe not. Here are two seemingly powerful arguments to the conclusion that it is rationally permissible for me to take the left, unsweetened box.

**Argument 3. I know that I have no preference for the contents of the right box**

Being rational involves, at least in part, acting on preferences between outcomes. So, surely:

*Recognition:*

Whenever I have two options, and I know that I have no preference between the outcome of the one and the outcome of the other, it is rationally permissible for me to take either.

4 'But they are not exactly the same', you might say. 'In the first case, the prospect associated with my taking the left box is  $\{<0.5, \text{I get A and could have gotten B+}>, <0.5, \text{I get B and could have gotten A+}>\}$ . In the second case the prospect associated with my taking the box alone is  $\{<0.5, \text{I get A and could have gotten A+}>, <0.5, \text{I get B and could have gotten B+}>\}$ . Different prospects.' True, and if, in addition to caring about what I get, I care about whether what I get is preferable to what I leave on the table, then I have reason to treat this difference as significant. But if I don't care about whether what I get is preferable to what I leave on the table, then I have no reason to treat this difference as significant.

In this case, I know that I have no preference between the outcome of my taking the left box and the outcome of my taking the right box. So it is rationally permissible for me to take the left box.

#### Argument 4: It is OK to defer to my better-informed self

Roughly: I know for sure that, if I were to see inside the boxes, I would have no preference for taking the right box. And it is rationally permissible for me to defer to my better-informed self.

More carefully: thinking of a *state of affairs* as a way for things to be, and thinking of a *maximal state of affairs* as a precise way for everything to be, here are two very plausible principles concerning rational permissibility:

##### *Deference*

If I know that any fully informed, rational person, with all and only my preferences between maximal states of affairs, would have a certain array of preferences between sub-maximal states of affairs on my behalf, then it is rationally permissible for me to have that array of preferences between sub-maximal states of affairs.

##### *Permissibility of action follows permissibility of preference*

If I have just two options, and it is rationally permissible for me to have no preference for my taking the one, and no preference for my taking the other, then it is rationally permissible for me to take the one and rationally permissible for me to take the other.

In this case I know that any fully informed, rational person, with all and only my preferences between maximal states of affairs, would have no preference for my walking away with the right box. So, by *Deference*, it is rationally permissible for me to have no preference for walking away with the right box. So, by *Permissibility of Action Follows Permissibility of Preference*, it is rationally permissible for me to walk away with the left box.

## 2. Two theories of decision

Suppose we make it our business to construct a general theory of rational decision under conditions of uncertainty that will tell us what it is rationally permissible to do when preferences are negatively intransitive. How might we accommodate each of these rival views? Accommodating the view that it is impermissible to take the left box is quite easy. Here's a rough statement of a theory I will call *prospectism*: we say that it is rationally permissible for me to take an action if and only if, for some way of rendering my preferences negatively transitive by keeping the preferences I have and adding new ones, the standard theory says that no alternative has higher expected utility.

Here's a more accurate, formal statement of the theory. First some terminology: where  $u$  is a function from outcomes to numbers, let the *prospect-extension* of  $u$ ,  $u^*$ , be a function from outcomes and prospects to numbers, such that for all outcomes  $o$ ,  $u^*(o) = u(o)$  and for all prospects  $p$ ,  $u^*(p) = \sum_o (Pp(o) \cdot u(o))$ , where  $o$  is a variable ranging over outcomes and ' $Pp(o)$ ' refers to the probability assigned to outcome  $o$  by prospect  $p$ . Say that  $u$  represents a coherent completion of my preferences between outcomes and prospects when, for all outcomes or prospects  $x$ ,  $y$ , if I prefer  $x$  to  $y$ , then  $u^*(x) > u^*(y)$ .

Now for the central claim:

### *Prospectism*

It is permissible for me to choose an option iff, for some utility function that represents a coherent completion of my preferences,  $u$ , no alternative has greater expected  $u$ -utility.<sup>5</sup>

In our original two opaque boxes case, prospectism says that it is rationally impermissible for me to take the left, unsweetened box. Why? Well, I prefer  $A+$  to  $A$ , and  $B+$  to  $B$ , so for any function,  $u$ , that represents a coherent completion of my preferences,  $u(A+) > u(A)$ , and  $u(B+) > u(B)$ . So for any function,  $u$ , that represents a coherent completion of my preferences, the expected  $u$ -utility of my taking the right box ( $0.5u(B+) + 0.5u(A+)$ ) is greater than the expected  $u$ -utility of my taking the left box ( $0.5u(A) + 0.5u(B)$ ).

Accommodating the view that it is permissible to defer to my known preferences between outcomes, to take the left, unsweetened box in the two opaque boxes case, is a little trickier.

A natural first move is to partition logical space into a set of *dependency hypotheses* – thinking of a dependency hypothesis as a maximally specific proposition concerning how things that matter to me causally depend on my present actions.<sup>6</sup> As a notational matter, let ' $P(d)$ ' refer to my credence that dependency hypothesis  $d$  is true, and let ' $od$ ' refer to the outcome of my

5 What is the 'expected  $u$ -utility' of an act? You can interpret this in different ways, depending on how you feel about Newcomb problems, causal and evidential decision theory. If you wish to be a causalist prospectist, then interpret it in a standard causalist way: the expected  $u$ -utility of act  $a$  is  $\sum_d (P(d) \cdot u(ad))$ , where  $d$  is a variable ranging over dependency hypotheses, propositions concerning how things beyond my control are, and ' $P(d)$ ' refers to my credence that hypothesis  $d$  is true, and ' $ad$ ' refers to the outcome of my taking act  $a$ , if hypothesis  $d$  is true. If you wish to be an evidentialist prospectist, then interpret it in a standard evidentialist way:  $\sum_d (P(d/a) \cdot u(ad))$ . This will not make a difference in any of the cases I discuss here.

6 This term was introduced by David Lewis in Lewis 1981. Other philosophers, and many decision theorists, talk of 'states' and 'states of nature'.

taking option  $o$ , if dependency hypothesis  $d$  is true. We can then restate the ‘Recognition’ principle from the last section in a more precise way:

*Recognition*

It is rationally permissible for me to choose option  $o$  if, for all alternatives open to me  $a$ , and all dependency hypotheses in which I have positive credence  $d$ , I do not prefer  $ad$  to  $od$ .

This is fine, so far as it goes, but it is not a general theory of decision, a theory that gives us full necessary and sufficient conditions for rational permissibility. It gives us one sufficient condition for rational permissibility, but it tells us nothing about cases in which the condition does not apply. For example:

*Two more opaque boxes*

You show me items A and B, a coin, and two opaque boxes. Then you toss the coin and, governed by the toss, place item A in one box and item B in the other. I don’t see which item went where. Then, with some determinate probability, you either do or do not switch the item in the right box with item C – *where C is an item that I prefer to both A and B*. I don’t see whether you made the switch.

In this case I do not know that I have no preference for the outcome of my taking the right box. I know that I have no preference for the outcome of my taking the *left* box. But maybe the right box contains item C. If it does, then I have a preference for the outcome of my taking the right box. The Recognition principle is silent.

What do we want our general theory to say about this sort of case? I suggest that we want it to say the following:

*Mild chancy sweetening*

When I do not have a strong preference for C over A and B, and my credence that you made the switch is small, it is rationally permissible to take the left box. (The right box has been *mildly sweetened* – not by a certain dollar, as in the original case, but by a small chance that it contains something that I regard as a little bit better than either A or B.)

Why? Well, we don’t want the theory to say that it is permissible to ignore a certain-dollar-sweetening, but impermissible to ignore (e.g.) a one-in-a-million-chance-of-a-hundred-dollars-sweetening. I far prefer a certain dollar to a one in a million chance of a hundred dollars.

*Powerful chancy sweetening*

When I have a strong preference for C over A and B, and my credence that you made the switch is large, it is rationally impermissible to take the left box. (The right box has been *powerfully sweetened* – by a large chance that it contains something that I regard as much better than A or B.)



Why? Well, obviously, if I am almost certain that the right box contains C, and C is a billion dollars, then I ought to take it.

Is there a moderately natural, general theory of decision that says all these things? I think so. Here's a rough statement of the theory I will call *deferentialism*: to find out if an action is permissible, I go to each relevant dependency hypothesis in turn. I take the coherent completion of my preferences that is most flattering to the action, supposing that the dependency hypothesis is true. I assign utilities to each of the actions open to me accordingly. I multiply these utilities by my credence that the dependency hypothesis is true ... and move on to the next dependency hypothesis. I sum up. If there is some way of doing this on which the action comes out ahead of (or at least *not behind*) its competitors, then the action is permissible.

Here's a more accurate, formal statement of the theory. First some terminology: let  $U$  be the set of utility functions that represent coherent completions of my preferences. Let a *regimentation* of  $U$  be a subset,  $R$ , of  $U$  such that for some outcomes  $A, B$ , for any function  $g$ ,  $g \in R$  iff  $g(A) = 1$  and  $g(B) = 0$ . (Note that it follows from the fact that utility functions are unique under positive affine transformation<sup>7</sup> that if  $R$  is a regimentation of  $U$ , then for each coherent completion of my preferences,  $R$  has one<sup>8</sup> and only one<sup>9</sup> representative of it as a member.)

Some more terminology: for any regimentation  $R$ , let the *dependency-expansion* of  $R$  be the set of functions  $f$  such that for any dependency hypotheses  $d$ , for some function  $r$  in  $R$ , for all actions open to me,  $a$ ,  $f(ad) = r(ad)$ . (Each function in the dependency-expansion of  $R$ , for each dependency hypothesis, agrees with some function in  $R$  on the utilities of the states of affairs that will come about if you act one way or another and the dependency hypothesis is true.)

Now for the central claim:

### *Deferentialism*

It is permissible for me to choose an option iff, for some regimentation,  $R$ , of the set of utility functions that represent my preferences, for some

7 Function  $v$  represents the same complete, coherent preferences as function  $u$  iff for some number  $i > 0$ , for some number  $j$ , for all  $x$ ,  $v(x) = iu(x) + j$ .

8 Proof: Take any coherent completion of my preferences, and a function that represents it,  $g$ . Now let  $h$  be the function such that for all  $x$   $h(x) = (1/g(A) - g(B)) \cdot g(x) - (g(B)/(g(A) - g(B)))$ . By construction,  $h(A) = 1$ ,  $h(B) = 0$ . And  $h$  represents the same coherent completion of my preferences as  $g$ , because for some number  $i > 0$ , for some number  $j$ , for all outcomes  $x$ ,  $g(x) = ih(x) + j$ .

9 Proof: Suppose that  $g$  and  $h$  are functions in  $R$  that represent the same coherent completion of my preferences. Because  $g$  and  $h$  are functions in  $R$ ,  $g(A) = h(A) = 1$ , and  $g(B) = h(B) = 0$ . Because  $g$  and  $h$  represent the same coherent completion of my preferences, for some number  $i > 0$ , for some number  $j$ , for all outcomes  $x$ ,  $g(x) = ih(x) + j$ . Solving for  $i$  and  $j$ ,  $i = 1$  and  $j = 0$ . Function  $g$  is function  $h$ .

function  $r$  in the dependency-expansion of  $R$ , no alternative has higher expected  $r$ -utility.<sup>10</sup>

In our original two opaque boxes case, deferentialism says that it is rationally permissible to take the left, unsweetened box. In that case there are two relevant dependency hypotheses. According to the first, the left box contains  $A$ , the right  $B+$ . According to the second, the left box contains  $B$ , the right  $A+$ . For any regimentation,  $R$ , of the utility functions that represent coherent completions of my preferences, some functions in  $R$  assign  $A$  utility greater than or equal to  $B+$ , others assign  $B$  utility greater than or equal to  $A+$ . Choose one of the former, call it  $u_1$ , and one of the latter, call it  $u_2$ . Notice that  $(0.5(u_1(A)) + 0.5(u_2(B))) \geq (0.5(u_1(B+)) + 0.5(u_2(A+)))$ , so for some function  $f$  in the dependency-expansion of  $R$ ,  $(0.5(f(A)) + 0.5(f(B))) \geq (0.5(f(B+)) + 0.5(f(A+)))$ , so for some function  $f$  in the dependency-expansion of  $R$ , the expected  $f$ -utility of taking the left box is greater than or equal to the expected  $f$ -utility of taking the right box.

And, more generally, deferentialism entails the principles I have called Recognition,<sup>11</sup> Mild Chancy Sweetening, and Strong Chancy Sweetening.

### 3. Which theory is right?

So much for prospectism and deferentialism. Which of them is right? I think this is a difficult, open problem. I feel the pull of both theories. But, on balance, I lean towards prospectism. It is not that I have a dazzling, decisive argument that goes significantly beyond what I have said already. It is rather that I am particularly moved by the thought that the excellent idea behind the standard theory of rational decision under conditions of uncertainty is to reduce a choice between items all of whose features you are not in a position to know (the outcomes of the options open to you) to a choice between items all of whose features you are in a position to know (supposing that you are in a position to know your own epistemic attitudes). But the only candidates I see for items all of whose features you are in a position to know

10 Again, please feel free to interpret 'expected  $r$ -utility' in the causalist or evidentialist way, depending on your feelings about Newcomb problems.

11 To be accurate: If we plug the causalist expected-utility formula into deferentialism then it entails the Recognition principle simpliciter. If we plug the evidentialist expected-utility formula into deferentialism then it entails the Recognition principle except in Newcomb cases – cases where I care about how things out of my control are, and my conditional credence in things out of my control being one way or another varies with actions available to me.

are the prospects associated with the options open to you. And thinking of the decision as a choice between prospects ties us to prospectism. Take the sugar.<sup>12</sup>

*The Massachusetts Institute of Technology*  
MA 02139-4307, USA  
casparh@mit.edu

## References

- Aumann, R.J. 1962. Utility theory without the completeness axiom. *Econometrica* 30: 445–62.
- Broome, J. 1997. Is incommensurability vagueness? In *Incommensurability, Incomparability and Practical Reason*, ed. Ruth Chang, 67–89. Princeton: Princeton University Press.
- Broome, J. 2000. Incommensurable values. In *Well-Being and Morality: Essays in Honour of James Griffin*, eds Roger Crisp and Brad Hooker, 21–38. Oxford: Clarendon Press.
- Chang, R. 2003. The possibility of parity. *Ethics* 112: 659–88.
- Chang, R. 2005. Parity, interval value and choice. *Ethics* 115: 331–50.
- Griffin, J. 1986. *Well-Being: Its Meaning, Measurement and Importance*. Oxford: Clarendon Press.
- Lewis, D. 1981. Causal decision theory. *Australasian Journal of Philosophy* 59: 5–30.
- Parfit, D. 1987. *Reasons and Persons*, paperback edition with corrections. Oxford: Oxford University Press.

12 Many thanks to Bob Stalnaker, Miriam Schoenfield, Brian Hedden, Nick Beckstead, Agustin Rayo, Steve Yablo and John Broome for most helpful discussions and comments.

## *Anti-Perfectionisms and autonomy*

BEN COLBURN

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