

TWO KINDS OF CONSEQUENTIALISM

Michael Smith
Princeton University

In a famous footnote in *Anarchy, State, and Utopia*, Robert Nozick wonders whether “a side-constraint view can be put in the form of the goal-without-side-constraint view” (Nozick 1974, 29). Put in more familiar terms, Nozick wants to know whether it is possible to represent side-constraints in a broadly *consequentialist* framework. For those who think that the arguments in favour of some sort of consequentialism are overwhelming, this question is of the first importance.

Nozick himself thinks that the issue is a tricky one, especially if the consequentialist is allowed to distinguish between two goals that an agent might have: the goal of there being no violation of one person’s rights *by another*—where I assume he means us to understand this in terms of a list of prohibited actions—and the goal of there being no violation of anyone’s rights *by the agent himself*. Indeed, the consequentialist might also need to distinguish between the goal of the agent himself not violating anyone’s rights *at any time* and the goal of the agent himself not violating anyone’s rights *at the time at which he acts*. But even if he is allowed to distinguish between such goals, Nozick thinks that the representation would still be objectionable because “gimmicky” (Nozick 1974, 29).¹

In recent years several theorists have addressed a more general version of Nozick’s question (Sen 1982, Vallentyne 1988, Broome 1991, Dreier 1993). Can all substantive moral theories, and perhaps all theories of reasons for action, both moral and non-moral, be represented in consequentialist terms? The arguments for and against the possibility of such a representation are complicated, but even if they could, Nozick’s residual objection that such representations are gimmicky remains. So let’s confront that question head-on. Are such representations gimmicky?

My own view is that they are not. Such representations of moral theories, and of theories of reasons for action more generally, are compulsory on

certain widely held and plausible views about the nature of reasons for action and value. As it happens, I hold one of these views myself (Smith 1994, Smith 2003, Smith 2005), but defending the correctness of some particular version such a view is strictly speaking unnecessary to answer the charge of gimmickiness. It suffices that the representation is compulsory on views that are both widely held and plausible. Since neither consequentialists nor those who hold these views about the nature of reasons for action and value seem to have noticed the connection between them, my main aim in this paper is to explain why I think that they are connected in this way.

1. What is Consequentialism?

Let's begin by getting clearer about consequentialism itself. Suppose an agent has certain options. Once we have fully specified the way the world would be if the agent were to take each of these options, and the value of each such world relative to those in which he takes other options—is the world better, or worse, or equal in value, or roughly equal in value, or incommensurable?—consequentialism tells us that we have thereby settled the deontic status of each of the agent's options: settled, in other words, whether the options are obligatory, permissible, forbidden, or indeterminate between these.

This is because consequentialism, at least as I understand it, is a *reductionist* doctrine. It reduces the deontic facts about an agent's actions to facts about the options she had when she performed that action and the value of the possible worlds in which she takes those options. Non-consequentialism, by contrast, is the denial of such a reduction. Consequentialism's core claim thus turns out to be about the *metaphysics of normativity*: normative facts of one sort, facts about the deontic status of actions, are reducible to normative facts of a quite different sort, facts about the values of the possible worlds in which those actions and their alternatives are performed.²

Moreover, so understood, consequentialism makes no claim beyond this: no claim about the nature of evaluative facts, no claim about what the value-making features are of the possible worlds that give rise to rankings, no claim about the completeness or otherwise of rankings, and no claim about the nature of the function from rankings to obligations. This is not to say that these are unimportant questions. It is simply to insist that, at least according to consequentialism, they are best understood as arising *in-house*. Though these and other important questions can be raised, they arise against a background of agreement on the core claim that defines consequentialism, that core claim being that we can reduce the deontic to the evaluative.

With this understanding of consequentialism in mind, I want to focus on the ways in which the answers that we give to two of these in-house questions interact with each other. The first is the question about the nature

of evaluative facts. Whereas some versions of consequentialism hold that for a consequence to have value is for an *evaluator-relative* fact about that consequence to obtain, other versions deny this, holding instead that for a consequence to have value is for an *non-evaluator-relative* fact about that consequence to obtain (Sen 1982, Regan 1983, Sen 1988—this will become clearer in what follows). The second is the question about the nature of the features that are value-making. Whereas some versions of consequentialism hold that the features that are value-making are one and all *neutral*, other versions deny this, holding instead that some or all of the value-making features are *relative* (Nagel 1970, Parfit 1984, Nagel 1986—this too will become clearer in what follows).

When Nozick wonders whether consequentialism can distinguish between the goal of there being no violation of one person's rights by another and the goal of the agent himself not violating anyone's rights, he is wondering about the latter question. An agent's action's having the feature being a non-violation *by himself* of anyone's rights *at the time at which he acts* is a relative value-making feature, rather than neutral, because in specifying the value-making feature we need to mention the agent himself and his time of action. Mention of one or another such feature is what it is for a value-making feature to be relative. (We will return to this point at the end.) Nozick's suggestion, on the consequentialist's behalf, is thus that he can represent side-constraints only if he allows that some value-making features are relative, rather than neutral. My suggestion, on the consequentialist's behalf, is going to be that this is in turn possible only if he further conceives of evaluative facts as evaluator-relative, rather than non-evaluator-relative. This is why the answers we give to the two questions interact.

The full set of options we have when we combine our answers to these two questions can be represented in the following matrix, or so I conjecture:

	ALL VALUE-MAKING FEATURES ARE NEUTRAL	NOT ALL VALUE- MAKING FEATURES ARE NEUTRAL
EVALUATIVE FACTS ARE NON-EVALUATOR- RELATIVE	NECESSARY	IMPOSSIBLE
EVALUATIVE FACTS ARE EVALUATOR-RELATIVE	POSSIBLE	POSSIBLE

If evaluative facts are non-evaluator-relative, then all value-making features are neutral, whereas if evaluative facts are evaluator-relative, then it is an open question whether value-making features are neutral or relative. The most fundamental issue to be addressed in attempting to answer

both Nozick's question about the relationship between side-constraints and consequentialism, and the more general question about the relationship between reasons for action as such and consequentialism, is thus whether evaluative facts are or are not evaluator-relative. This is because the stand we take on this issue fixes whether consequentialism allows for the existence of relative value-making features, or whether it restricts us to supposing that value-making features are one and all neutral.

The remainder of the paper comprises four further sections. In the next section I consider the standard argument for supposing that consequentialism entails that all value-making features are neutral and I show that this argument does indeed have, as a crucial premise, the claim that evaluative facts are non-evaluator-relative. In the third section I show that if we suppose instead that evaluative facts are evaluator-relative, then that argument collapses. In the fourth section I show how widely shared and plausible the evaluator-relative conception of evaluative facts is and I explain why those who have such a conception should accept some form of consequentialism. As we will see, this suffices to put to rest Nozick's charge that the consequentialist representation of side-constraints is gimmicky. And then in the fifth and final section I respond to one line of objection to what's been said.

2. The Standard Argument for Supposing that all Value-making Features are Neutral Rather than Relative and why Some Think that this Augurs Against Consequentialism

Consider once again Nozick's famous footnote. Suppose we allow the consequentialist to distinguish between the goal of there being no violation of one person's rights by another at any time and the goal of an agent's not violating anyone's rights himself at the time at which he acts. Would this suffice for the consequentialist to represent side-constraints?

The answer is that it would not, at least not all by itself. The reason why is implicit in G. E. Moore's famous refutation of ethical egoism in *Principia Ethica* (1903). Imagine two people, A and B, and imagine further that A's action has the feature of not being a violation of anyone's rights by himself at the time at which he acts and that B's action has the feature of not being a violation of anyone's rights by himself, that is B, at the time at which he acts. Now think of the distribution of value. A's action has a value-making feature and so does B's. But if A's and B's actions both have value-making features, then how could the value-making feature possessed by B's act, along with that possessed by A's, fail to affect the value of the possible worlds in which A acts, if A's act has B's act as a consequence? To think that it could fail to affect the value of those possible worlds is to suppose, implausibly, that only some, rather than all, of the value-making features present in a possible world affect the value of that world.

Suppose, for example, that A has the option of not violating anyone's rights himself at the time at which he acts or the option of bringing it about that B does not violate anyone's rights himself at the time at which he acts. These options are indistinguishable in terms of value-making features, so it seems that the possible worlds in which A performs each of these options must have the same value, other things being equal. Or suppose that A has the option of bringing it about that he does not violate anyone's rights himself at the time at which he acts, but as a consequence B performs an act that does violate someone's rights at the time at which he acts, and that A also has the option of violating someone's rights himself at the time at which he acts, but, as a consequence, B does not violate anyone's rights at the time at which he acts. Once again, since the possible worlds in which these options are taken have exactly the same value-making features, it seems that they too must be indistinguishable in terms of value, at least other things being equal.

If this is agreed then it follows immediately that simply distinguishing between different value-making features in the manner suggested by Nozick isn't enough to allow for the representation of side-constraints in consequentialist terms. For side-constraints are obligations agents have not to violate anyone's rights themselves at the time at which they act. If this view is to be represented in consequentialist terms, there must therefore be some difference in the value of the possible worlds in which agents take this option over the possible worlds in which they take some alternative. But in the possible worlds just described, there is no difference in value.

Some theorists think it follows from this that side-constraints cannot be represented in consequentialist terms at all (Broad 1942). If there are side-constraints, these theorists suggest, then we must admit that there is at least one deontic fact, namely, that captured by:

SIDE-CONSTRAINT: (x)(t)(x at t ought not to bring it about that he violates anyone's rights himself at t)

whose truth does not supervene on evaluative facts. But I don't think that we should leap to this conclusion just yet. There is a deeper explanation of why merely distinguishing between different value-making features in the manner suggested by Nozick isn't enough to allow us to represent side-constraints within a consequentialist framework, an explanation which leaves intact the possibility that we can represent side-constraints in consequentialist terms.

In order to see what this deeper explanation is, let's unpack the claim that when A does not violate anyone's rights himself at the time at which he acts, and when B does not violate anyone's rights himself at the time at which he acts, both of their actions have a value-making feature. What is it about these features of their acts that makes them value-making features? The answer to this question has to be that they are both features in virtue of which

their acts have value. But what exactly does this mean? In particular, does it mean that these are features in virtue of which actions have the *same* feature or *different* features? This is where the distinction between an evaluator-relative conception of evaluative facts, as opposed to a non-evaluator-relative conception, comes to the fore.

Think for a moment about how G. E. Moore would answer this question. In virtue of not being a violation of anyone's rights by A at the time at which he acts, A's act has the simple feature of being good, and in virtue of not being a violation of anyone's rights by B at the time at which B acts, B's act also has the simple feature of being good. In virtue of having their respective value-making features, A's act therefore has the *very same feature* as that possessed by Bs: the goodness that A's and B's acts each have is a simple feature, and hence not itself relative to either A or B. This Moorean conception of goodness (and badness) as a simple feature entails a non-evaluator-relative conception of evaluative facts par excellence. An evaluative fact obtains, according to Moore, just in case something possesses the simple feature of goodness (or badness). Now suppose that A violates someone's rights and, as a consequence, B does not violate anyone's rights. Though some badness is located in A's act, some goodness is located in B's. Or suppose that A does not violate anyone's rights and, as a consequence, B does violate someone's rights. Then though some badness is located in B's act, some goodness is located in A's act. The possible worlds in which A performs these two acts must have the same value, other things being equal, because goodness and badness are instantiated in those worlds in exactly the same measure.

What's remarkable here is that, notwithstanding the fact that we initially characterized the value-making features of A's and B's acts in relative terms, the effect of supposing that the value possessed by acts with these value-making features is itself a simple feature, à la Moore, is to ensure that the contribution that these value-making features make is exactly the same as the contribution made by their neutral counterparts. Thus, for example, once we've fully understood the way in which a value-making feature like an agent's not violating anyone's rights himself makes for Moorean value, we see that all that can really make a contribution to Moorean value is his acting so as to bring it about that there is no violation of someone's rights by someone. The reference made to the agent himself in our initial characterization of the value-making feature is otiose.

The initial conclusion we must draw is thus that, if evaluative facts are non-evaluator-relative *in the manner of Moore*, then all value-making features are neutral. But my conjecture is that this same conclusion about the neutral nature of value-making features follows on any conception of evaluative facts as non-evaluator-relative. That same conclusion follows because the argument just given didn't turn on anything special about Moore's own version of the non-evaluator relative view. The crucial point is simply that, even if we do initially characterize value-making features in relative terms,

on a non-evaluator-relative conception of evaluative facts, there will be no way to prevent the value-making features possessed by actions performed in a possible world, no matter who performs them or when they perform them, from making a contribution to the evaluative facts about that possible world that are relevant to the deontic status of agents' acts, no matter who those agents are or when they act. The only way to prevent this is for the evaluative facts themselves to be in some way relativized. The upshot is that, on a non-evaluator-relative conception of evaluative facts, the real contribution to value made by relative value-making features is that made by the value when neutrally characterized.

3. Why Consequentialism is Consistent with Value-making Features being both Neutral and Relative even so: Evaluator-relative vs Non-evaluator-relative Conceptions of Evaluative Facts

Note how different things look if we conceive of evaluative facts in evaluator-relative terms. Imagine, for example, that we conceive of goodness not in the manner of Moore, but in the manner of Henry Sidgwick (1907). An outcome is good, according to Sidgwick, just in case there is "a rational dictate to aim at [it]" (Sidgwick 1907, 112). This account of goodness entails an evaluator-relative conception of evaluative facts par excellence because, for evaluative facts to obtain on such a view, the value-making features of states of affairs must bear the right kind of relation to the aims or preferences of those who are subject to rational dictates.

Suppose, for example, that the following is a basic norm of rationality:

(x)(t)(It is a rational dictate that x at t prefers that there be no violation of anyone's rights by x at t)

In that case, in virtue of not violating anyone's rights himself at the time at which he acts (that is to say, in virtue of having a relative value-making feature), A's act is one that there is a dictate of reason that A prefers (that is to say, A's act has value_A), and in virtue of not violating anyone's rights himself at the time at which he acts (that is to say, in virtue of having a relative value-making feature), B's act is also one that there is a dictate of reason that B prefers (that is to say, B's act has value_B). There is thus no value that A's and B's acts both have in virtue of having their respective relative value-making features.

If we conceive of evaluative facts in this evaluator-relative way, then note that the Moorean argument considered in the previous section collapses. If the only dictate of reason is the one just formulated, and if A has the option of not violating anyone's rights himself at the time at which he acts, and, as a consequence, B violates someone's rights at the time at which he acts, and A also has the option of violating someone's rights himself at the time at

which he acts, but, as a consequence, B does not violate anyone's rights at the time at which he acts, then A's preferences will violate a rational dictate if he doesn't prefer the former to the latter. Likewise for B. This is because there is no rational dictate telling A or B to aim at *each other's* not violating anyone's rights. There is, in other words, no way to derive the value_B of an outcome from its value_A, or vice versa.

This in fact provides us with a solution to a puzzle about evaluator-relative conceptions of evaluative facts. When Amartya Sen originally put forward the idea that we should conceive of evaluative facts in evaluator-relative terms, Donald Regan wondered why, in a situation like that just described, A's obligations would have to be a function of the value_A of his actions, rather than the value_B of his actions (Regan 1983). Why shouldn't we suppose instead that A's obligations are a function of *both* the value_A *and* the value_B of his actions, or perhaps the value_{someone} of his actions? To suppose that A's obligations have to be a function of the value_A of his actions is to make an arbitrary distinction among the values which give rise to obligations, or so Regan suggested. In this way he hoped to reinstate Moore's original argument, but in the context of an evaluator-relative conception of values.

But the answer to Regan's question should now be obvious—or rather, it should be obvious, given that our route to an evaluator-relative conception of evaluative facts has been via a Sidgwickian conception of value. For on that conception, A's obligations have to be a function of the value_A of A's actions, rather than the value_B, because this is the only way in which we can guarantee that A's obligations track outcomes that there is a rational dictate that A prefers. Likewise for B. In other words, A's obligations have to be a function of the value_A of A's actions because only so can we guarantee the *normativity* of obligation.

We are finally in a position to see what the consequentialist must do in order to formulate side-constraints in consequentialist terms. First, he must distinguish between the neutral goal of there being no violation of one person's rights by another at any time and the relative goal of an agent's not violating anyone's rights himself at the time at which he acts, and he must suppose that the value-making feature of an agent's actions that is relevant for the purposes of formulating side-constraints is the one specified by the relative goal, rather than the neutral goal: that is, he must focus on the value that actions have in virtue of their being non-violations of anyone's rights by the agent himself at the time at which he acts. So far we are in agreement with Nozick.

Second, however, and just as importantly, the consequentialist must conceive of the value that outcomes have in virtue of having value-making features in evaluator-relative terms. If what it is for something to be of value is for (say) there to be a rational dictate that agents prefer that thing, and if the only rational dictate is that each agent must prefer that there be no violations of anyone's rights by himself at the time at which he acts, then

consequentialism suffices to explain the truth of SIDE-CONSTRAINT. For in that case the deontic facts captured by SIDE-CONSTRAINT supervene on the relative value-making features of acts and the evaluator-relative evaluative facts.

Having said this, however, it is important to note that buying into an evaluator-relative conception of evaluative facts doesn't itself *entail* that all value-making features are themselves relative. It is consistent with the evaluator-relative conception of evaluative facts that some value-making features are neutral. Suppose, for example, that:

(x)(t)(It is a rational dictate that x at t prefers that there be no violations of anyone's rights by anyone to there being some, and that there be fewer violations to there being more)

In that case, actions would have value in virtue of their having the neutral feature of bringing about as few violations of anyone's rights as possible. If this were the only rational dictate concerning people's preferences then consequentialists would have to reject SIDE-CONSTRAINT in favour of:

CONSEQUENTIALISM OF RIGHTS: (x)(t)(x at t ought to bring it about that there are as few violations of anyone's rights by anyone as possible)

For the truth of CONSEQUENTIALISM OF RIGHTS would supervene on the neutral value-making features that actions have together with the evaluator-relative evaluative facts.

Nor does buying into an evaluator-relative conception of evaluative facts itself entail that value-making features are themselves either all relative or all neutral. It is consistent with the evaluator-relative conception of evaluative facts that some value-making features are neutral and some are relative. On the Sidwickian conception, this will be so if there are two rational dictates, one telling people to prefer some state of affairs characterized in neutral terms and one telling them to prefer some state of affairs characterized in relative terms. This would, of course, require us to conceive of the preferences in question as preferences other things being equal, rather than all things considered. But I take it that there is no reason to suppose either that there are no such preferences or that there are no rational dictates concerning them.

Nor again does buying into an evaluator-relative conception of evaluative facts entail that, if there are both neutral and relative value-making features, then there is incommensurability. It is consistent with the evaluator-relative conception of evaluative facts that neutral and relative value-making features are commensurable, roughly commensurable, and incommensurable. On the Sidwickian conception of value, it all depends on whether there is a rational dictate telling us how to combine the two orderings of preferences other

things being equal into an ordering of preference all things considered. I take it that there is no reason to deny the existence of such a rational dictate (see Smith 2003 and Parfit forthcoming, but contrast Sidgwick 1907 and Smith 2009).

Let's recap. I said at the outset that in order to capture a side-constraint view it isn't enough for a consequentialist to distinguish between relative and neutral value-making features. We have now seen why this is so. If the consequentialist has a non-evaluator-relative conception of evaluative facts, then the contribution that relative value-making features make to value will be the same as that made by their neutral counterparts: the distinction between the two sorts of value-making features will be a distinction without a difference. To capture a side-constraint view the consequentialist must therefore, in addition, reject a non-evaluator-relative conception of evaluative facts in favour of an evaluator-relative conception of some sort or other.

Having arrived at this point the question we must ask next is whether this suffices to make the consequentialist representation of side-constraints non-gimmicky. To anticipate, I will argue that it not only suffices to make the representation non-gimmicky, but that it makes the representation compulsory on a certain widely held and plausible account of the relationship between reasons for action and value.

4. Is the Proposed Representation of Side-constraints in Consequentialist Terms Gimmicky?

The question with which we began was whether the representation of side-constraints in consequentialist terms is gimmicky. Now that we've seen what a consequentialist would have to say in order to represent side-constraints, what should we make of this charge? The answer comes in two parts.

The first part is a reminder of why some people think that the arguments in favour of some form of consequentialism are overwhelming. As Joseph Raz points out, one of the dominant views theorists have held about reasons for action—he calls it the “classical view”—is that an agent has a reason to act in a certain way just in case something good will result from his acting in that way (Raz 1999, especially Chapters 11–13). Raz traces this view back to Plato and Aristotle, but it is also extremely widely held by contemporary theorists, both consequentialists (see, for example, Parfit forthcoming), and non-consequentialists alike (see, for example, Anscombe 1957, and of course Raz himself). But it is difficult to see how those who hold this view of reasons for action could give anything other than a consequentialist explanation of what agents ought to do.

In order to see that this is so, consider side-constraints. If those who accept the classical view are to resist a consequentialist explanation of

side-constraints of the kind given above, such theorists must suppose, on the one hand, that we each have an obligation not to violate people's rights ourselves, and yet also hold, on the other, either that we have no reason to act in that way, which is inconsistent with the normativity of obligation, or that we have a reason that doesn't entail the good of what will thereby come about, which is a denial of the classical view. On the face of it, this combination of views looks to be flat-out inconsistent. Those who hold the classical view thus seem to be committed to some form of consequentialism. Of course, many who hold the classical view don't see things this way themselves. But we are now in a position to hazard a guess as to why not.

Why doesn't everyone who holds the classical view see that their view commits them to consequentialism? My guess is that they hold the mistaken view that, according to consequentialism, all value-making features are neutral, and that the reason they think this is in turn that they think that consequentialism is committed to a non-evaluator-relative conception of evaluative facts. Unsurprisingly, they conclude that the reasons provided by side-constraints must be reasons of some special kind. But no special accommodation of the reasons provided by side-constraints is required if you instead have an evaluator-relative conception of evaluative facts and hold that some value-making features are themselves relative. The reasons for action provided by side-constraints are exactly the same as any other reasons.

The second part of the response to the charge of gimmickiness is to underscore both how widely accepted evaluator-relative conceptions of evaluative facts are and how strongly different versions of this conception all augur in favour of some form of consequentialism, despite the differences between them. For instance, as we've already seen, side-constraints are a natural upshot of the idea that there is a rational dictate telling each agent to aim at his own non-violation of anyone's rights. Indeed, once the existence of Sidwickian rational dictates concerning aims is agreed, it is difficult to imagine what apart from the fact that side-constraints answer to some such rational dictate the rationale for them could be. For why shouldn't we violate people's rights ourselves if no rational dictate bears on whether we should or shouldn't aim at our doing so? The existence of such an obligation, in the absence of such a rational dictate, would seem to rob the obligation of its normative force.

Or consider how matters look to those who eschew talk of rational dictates, but who hold that goodness is a matter of what it would be especially *fitting* for us to desire (Ewing 1947, Ewing 1959). They should similarly find it difficult to imagine what a non-consequentialist rationale for side-constraints would be. For suppose it isn't fitting for us to desire that we do not violate people's rights ourselves. In that case, why shouldn't we violate their rights? Or consider things from the perspective of those who eschew talk of rational dictates and fittingness, but who hold that goodness is a matter of what

merits approval (McDowell 1985, Wiggins 1987). They should surely wonder why we shouldn't violate people's rights ourselves if our violations don't merit disapproval. Similarly for those who hold that goodness is a matter of what we would want if we were fully rational (Smith 1994). They should wonder why we shouldn't violate people's rights ourselves if our fully rational counterparts wouldn't desire that we not violate their rights. And likewise for those who hold that goodness is a matter of what *there is reason* to desire (Scanlon 1998). For why shouldn't we violate people's rights ourselves if there is no reason not to desire that we do so?

The upshot should now be clear. It can hardly be *gimmicky* to represent side-constraints in consequentialist terms given that one of the dominant views about reasons for action is that agents have reasons to act in a certain way just in case there is some good to be realized by their acting in that way, and given that so many theorists accept a version of the evaluator-relative conception of evaluative facts which makes it not only possible, but seemingly compulsory, to represent the goods that side-constraints would have to promote if we were to represent them in consequentialist terms. I conclude that the charge of gimmickiness is therefore unjustified.

5. An Objection?

Or is it?

Mark Schroeder calls the sort of view I've been trying to defend the coherence of in this paper 'Agent Relative Teleology (ART)' (Schroeder 2007). Contrary to what I've just suggested, however, he thinks that when we develop ART alongside (say) a fitting-attitudes version of the evaluator-relative conception of evaluative facts, the resulting view is vulnerable to a serious line of objection. It is worth stating Schroeder's objection in full:

The problem that I have with this idea is simple. It is not that I am opposed to fitting-attitudes analyses of 'good'. It is just that from some very straightforward and compelling assumptions, it follows that the fitting-attitudes analyses of agent-neutral and agent-relative value cannot both be true—at least if we are to make the assumptions about what is better than what, and what is better-than-what-relative-to whom, which ART-ists need.

The assumptions that I need are that 'good' is a gradable adjective and that gradable adjectives need to be analyzed in terms of their comparative form. As with 'tall', 'taller', and 'tallest', and 'fat', 'fatter', and 'fattest', it is the better than relation that is basic—something is good when it is better than sufficiently many things (perhaps in a contextually relevant comparison class), and it is best when it is better than everything (in some comparison class). So from this assumption, it follows that if a fitting-attitudes analysis of good is correct, it must really be a fitting-attitudes analysis of better than. And it is easy to construct such an account—A is better than B just in case it is fitting to prefer A to B. Fitting

for whom? Well, for everyone. So to have the same structure as this account (which, after all, is the whole point of this idea), the fitting-attitudes analysis of agent-relative value would start with an account of better-than-relative-to. It would say that A is better-than B relative-to Jon just in case it is fitting for Jon to prefer A to B.

And now we have enough to generate a problem. For from these two accounts, it follows that A is better than B just in case A is better-than B relative-to everyone. For to be better than B is to be fittingly preferred by each, but being fittingly preferred by anyone is being better-relative-to her. So, in particular, it follows that if A is agent-neutrally better than B, then it is better-than B relative-to, say, Franz. But this conclusion makes it impossible to reconstruct what was supposed to be puzzling about constraints for ordinary consequentialism!

Constraints, recall, were cases like that in which Franz can murder to prevent Hans and Jens from murdering. They were supposed to be paradoxical, because it was supposed to be agreed on all sides that things are (agent-neutrally) worse if Franz does not murder—for then there are two murders, rather than one. Moreover, in order to accommodate constraints, ART-ists must suppose that things are better-relative-to Franz, if he does not murder. So in order to agree that there was a puzzle in the first place, and also have an answer to it, ART-ists need to suppose that constraints are cases in which it is agent-neutrally better for Franz to murder, but agent-relatively better for him not to murder. But from the fitting-attitudes analyses of both agent-relative and agent-neutral value, what we just derived was the thesis that if it is agent-neutrally better for Franz to murder than not murder, then it is better-relative-to Franz for Franz to murder than not murder. Which surely has to be inconsistent with its being better-relative-to Franz for Franz to not murder than to murder.

The foregoing argument shows that ART-ists cannot have it both ways. They have to give something up. (Schroeder 2007, 293–4)

This is a devastating objection if it is correct. But is it correct?

Let me restate Schroeder's objection. At least as I understand it, his objection is that when we fill out the evaluator-relative conception of evaluative facts in terms of a fitting-attitudes story, it turns out that we can derive evident falsehoods. In particular, he thinks it follows from this way of filling out the evaluator-relative conception of evaluative facts that if something is a neutral value-making feature then *that very feature* is also a relative value-making feature, and he further thinks, plausibly enough, that this is absurd. He concludes that something must therefore be wrong with filling out an evaluator-relative conception of evaluative facts in terms of a fitting-attitudes story.

There are many things to say about this objection, but perhaps the most obvious thing to say is that it conflates the relativity of relative value-making features with the relativity of evaluator-relative evaluative facts. To be sure, if we have an evaluator-relative conception of evaluative facts, as we do if we have a fitting-attitudes account of what it is for something to be valuable, then it follows that all evaluative facts are relative to evaluators. Thus, for example, if

there being no murders is better than there being some, and there being fewer is better than there being more, then it follows that it is fitting for everyone to prefer no murders to some, and fewer to more, and it further follows that it is fitting for Franz to have these preferences, and for Hans to have these preferences, and so on. In other words, according to the evaluator-relative conception, if something is better then it follows that it is better_{everyone}, and if something is better_{everyone} then it follows that it is better_{Franz}, better_{Hans}, and so on. This all follows if we hold an evaluator-relative conception of evaluative facts.

The crucial question to ask, however, is whether it follows from the evaluator-relativity of evaluative facts that the *value-making features* in question are themselves both neutral and relative. Schroeder thinks that it does follow because he thinks that what makes a value-making feature neutral is that it is better_{everyone} and that what makes a value-making feature relative is that better_{Franz}, better_{Hans}, and so on. But this is a mistake because, to repeat, the relativity that Schroeder here describes is the relativity of evaluator-relative evaluative facts, not the relativity of relative value-making features. As we saw earlier, the mere fact that evaluative facts are relative to evaluators tells us nothing about whether the value-making features themselves—that is to say, on the fitting attitudes story, the contents of the preferences that it is fitting to have—are themselves relative or neutral. This is why, given the evaluator-relative conception of evaluative facts, we could explain the truth of both SIDE-CONSTRAINTS and CONSEQUENTIALISM OF RIGHTS. Whether value-making features are themselves neutral or relative turns on *what* it is fitting to prefer, not the fact that value as such consists in fittingness to prefer.

If this is right then it must be possible to characterize the difference between relative and neutral value-making features in some quite different to that proposed by Schroeder. So how should we characterize this difference? The answer is I hope obvious. We should say that what makes a value-making feature relative is that, in characterizing the value-making feature itself—that is to say, on the fitting-attitudes account, in characterizing the content of the preference that it is fitting to have—we have no choice but to explicitly mention the evaluator to whom the evaluative facts are relative. In the situation described above this isn't so, for no explicit mention of (say) Franz was required in order to state the content of Franz's preference. What we said is that it is fitting for Franz to prefer no murders to some, and fewer to more. Ditto for Hans, and so on. The value-making feature in question is therefore, as expected, neutral.

Explicit mention is, however, required when it is better that each (say) not murder rather than murder. For, if we have a fitting-attitudes account, what makes this true is that it is fitting for each to prefer that he not murder rather than murder. Of course, it follows that it is fitting for Franz to prefer that Franz not murder rather than murder, and fitting for Hans to

prefer that Hans not murder rather than murder, and so on, so in this case too it follows that if something is better then it is better_{everyone}, and that if it is better_{everyone} then it is better_{Franz}, better_{Hans}, and so on. This is just the evaluator-relativity of evaluative facts again. But the crucial difference between this case and that described above is that, in this case, the thing that is better—that is, the content of each person's preference—is something that we can characterize only if we mention each. It is fitting for Franz to prefer that *Franz* not murder rather than murder, and fitting for Hans to prefer that *Hans* not murder rather than murder, not vice versa. The value-making feature is therefore relative, just as expected.

To sum up, Schroeder objects that when we fill out the evaluator-relative conception of evaluative facts in terms of a fitting-attitudes story, it turns out that we can derive evident falsehoods. But the evident falsehoods are only derivable because Schroeder conflates the relativity of relative value-making features with the relativity of evaluator-relative evaluative facts that I have been at pains to distinguish. Once we do distinguish between these, and once we properly characterize what, according to the fitting-attitudes story, it is for value-making features to be both neutral and relative, Schroeder's objection lapses.³

Notes

1. For the record, it is unclear to me that representing side-constraints in consequentialist terms requires us to suppose that there is value in the outcome of an agent's not violating anyone's rights himself. It would be enough for the agent's not violating anyone's rights to be a *condition* of the value of (some or all of) the outcomes of the actions he performs. When Nozick says that the representation of side-constraints as a valuable outcome is gimmicky, I sometimes wonder whether he was objecting to their representation as a valuable outcome, rather than as a condition of the value of outcomes that are valuable. Nozick also wonders whether, assuming that the consequentialist does represent side-constraints as a valuable outcome, he should assign the goal of not violating anyone's rights at the time at which he acts *infinite weight*. For an argument against supposing that side-constraints are absolute in the way they would have to be if they had infinite weight see Frank Jackson and my 'Absolutist Moral Theories and Uncertainty' (Jackson and Smith 2006). Nozick himself remains non-committal on whether side-constraints are absolute (Nozick 1974, 30).
2. If consequentialism is the view that deontic facts reduce to evaluative facts, does it follow that consequentialists are one and all *cognitivists* about moral judgement? No it doesn't. Non-cognitivists have their own way of understanding talk of normative facts, so I assume that they have their own way of understanding what a consequentialist has in mind when he affirms, or a non-consequentialist when he denies, that deontic facts reduce to evaluative facts. For more on whether non-cognitivists who are consequentialists are able to distinguish neutral from relative

value-making features see James Dreier's 'Accepting Agent Centred Norms' (Dreier 1996).

3. Earlier versions of this paper were delivered at Harvard University, the University of Missouri at Columbia, and the Australian National University. I am grateful to all those who participated in the discussions for their helpful comments.

References

- Anscombe, G. E. M. 1957: *Intention* (Oxford: Basil Blackwell).
- Broad, C. D. 1942: 'Certain Features of Moore's Ethical Doctrines' in P. A. Schilpp (ed.) *The Philosophy of G. E. Moore* (Evanston and Chicago, Northwestern University). 41–67.
- Broome, John 1991: *Weighing Goods* (Oxford: Blackwell).
- Dreier, James 1993: 'Structures of Normative Theories' *The Monist* (76) 22–40.
- 1996: 'Accepting Agent Centered Norms' *Australasian Journal of Philosophy* (74) 409–22.
- Ewing, A. C. 1947: *The Definition of Good* (London: MacMillan).
- Ewing, A. C. 1959: *Second Thoughts in Moral Philosophy* (London: Routledge and Kegan Paul).
- Jackson, Frank and Michael Smith 2006: 'Absolutist Moral Theories and Uncertainty' *Journal of Philosophy* (103) 267–28.
- McDowell, John 1985: 'Values and Secondary Qualities' in Ted Honderich (ed.) *Morality and Objectivity* (London: Routledge and Kegan Paul). 110–29.
- Moore, G. E. 1903: *Principia Ethica* (Cambridge: Cambridge University Press).
- Nagel, Thomas 1970: *The Possibility of Altruism* (Princeton: Princeton University Press).
- 1986: *The View from Nowhere* (Oxford: Oxford University Press).
- Nozick, Robert 1974: *Anarchy, State, and Utopia* (New York: Basic Books).
- Parfit, Derek 1984: *Reasons and Persons* (Oxford: Clarendon Press).
- forthcoming: *On What Matters* (Oxford: Clarendon Press).
- Raz, Joseph 1999: *Engaging Reason: On the Theory of Value and Action* (Oxford: Oxford University Press).
- Regan, Donald H. 1983: 'Against Evaluator Relativity: A Response to Sen' *Philosophy and Public Affairs* (12) 93–112.
- Scanlon, T. M. 1998: *What We Owe to Each Other* (Cambridge, MA: Harvard University Press).
- Schroeder, Mark 2007: 'Teleology, Agent-Relative Value, and "Good"' *Ethics* (117) 265–295.
- Sen, Amartya 1982: 'Rights and Agency' *Philosophy and Public Affairs* (11) 3–39.
- 1988: 'Evaluator Relativity and Consequential Evaluation' *Philosophy and Public Affairs* (12) 113–132.
- Sidgwick, Henry 1907: *The Methods of Ethics* (Indianapolis: Hackett Publishing Company, 1981).
- Smith, Michael 1994: *The Moral Problem* (Oxford: Basil Blackwell).
- 2003: 'Neutral and Relative Value after Moore' *Ethics* (113) 576–598.
- 2005: 'Metaethics' in Frank Jackson and Michael Smith (eds.) *Oxford Handbook of Contemporary Philosophy* (Oxford: Oxford University Press).
- 2009: 'Desires, Values, Reasons, and the Dualism of Practical Reason' *Ratio* (22).
- Vallentyne, Peter 1988: 'Gimmicky Representations of Moral Theories' *Metaphilosophy* (19) 253–263.
- Wiggins, David 1987: 'A Sensible Subjectivism' reprinted in his *Needs, Values, Truth* (Oxford: Basil Blackwell).