BEN BRADLEY

EXTRINSIC VALUE

(Received in revised form 19 November 1996)

THE PROBLEM

In *Republic* 357, Plato distinguishes three kinds of goods:

> "Tell me, is there in your opinion a kind of good that we would choose to have not because we desire its consequences, but because we delight in it for its own sake – such as enjoyment and all the pleasures which are harmless and leave no after effects other than the enjoyment in having them?"
>
> "In my opinion, at least," I said, "there is a good of this kind."
>
> "And what about this? Is there a kind we like both for its own sake and for what comes out of it, such as thinking and seeing and being healthy? Surely we delight in such things on both accounts."
>
> "Yes," I said.
>
> "And do you see a third form of good, which includes gymnastic exercise, medical treatment when sick as well as the practice of medicine, and the rest of the activities from which money is made? We would say that they are drudgery but beneficial to us; and we would not choose to have them for themselves but for the sake of the wages and whatever else comes out of them."
>
> "Yes, there is also this third," I said, "but what of it?"

The three kinds of goods Plato distinguishes are intrinsic goods, instrumental goods, and goods that are both intrinsic and instrumental. Many philosophers still consider this division of goods to be exhaustive; for example, John Hospers writes that

> Our question, "What things are good?" is ambiguous. There are some things which we consider good (or desirable) only for their results – for what they lead to. There are other things which we consider good not because of what they lead to but because of what they are in themselves . . . (Hospers 104).

Louis Pojman apparently agrees when he writes that "the essential difference [between kinds of goods] is between intrinsic and instrumental goods" (Pojman 85).[1] William Frankena makes a similar claim when he identifies instrumental value with extrinsic value:[2] "One may also say that something is good on the ground that it is

Philosophical Studies **91:** 109–126, 1998.

a *means*, necessary, sufficient, or both, to a good end, as when one says, 'It is a good idea to go to the dentist twice a year.' Then it is *extrinsically* or *instrumentally* good, or good *as a means*" (Frankena 65, emphasis his).

I believe these views are mistaken. It seems possible that something may have value that is neither intrinsic nor instrumental. For example, something could be good not because of what it causes or is a means to, but rather because of what it *signifies*.[3] Thus there seems to be at least one more kind of value: "signatory" value. Another type of extrinsic value, mentioned by C. I. Lewis (1955), is "contributory" value, or the value something has in virtue of being a part of a valuable whole. If signatory and contributory value are genuine types of value, they are obviously types of extrinsic value; so extrinsic and instrumental value cannot be identified. This leaves us with a problem: if extrinsic value is not instrumental value, then what is it?

Even more surprising is the fact that the notion of instrumental value remains misunderstood. Pojman claims that something is instrumentally good only if it is a means to something intrinsically good (Pojman 85); Richard Brandt claims that something is instrumentally good only if its expected results are intrinsically good (Brandt 303).[4] Hospers writes: "The first kind of good is called *instrumental* good because the goodness or worthwhileness of these things lies in their being instruments toward the attainment of the other things which are considered good not merely as instruments" (Hospers 104–105, emphasis his). This view of instrumental value is so widely held that Gilbert Harman includes it as part of the "standard account" of intrinsic and instrumental value (Harman 796).

Yet it has been noted that something might be instrumentally good in virtue of causally contributing to *preventing* something intrinsically *evil*; if this is so, then something could be instrumentally good without bringing about, or being expected to bring about, anything intrinsically good. This has been noticed recently by Earl Conee (1980), but seems to have been noticed long ago by Jeremy Bentham: "By utility is meant that property in any object, whereby it tends to produce benefit, advantage, pleasure, good, or happiness . . . or . . . *to prevent the happening of mischief, pain, evil, or unhappiness to the party whose interest is considered*" (Bentham 2, my emphasis).

J.S. Mill also seems to have recognized the existence of preventive value: "all desirable things . . . are desirable either for pleasure inherent in themselves or as means to the promotion of pleasure *and the prevention of pain*" (Mill 10–11, my emphasis).

It is my contention that any correct analysis of the concept of instrumental value must account for prevention value, and that any correct analysis of extrinsic value must account not only for instrumental (hence preventive) value, but also for signatory and contributory value. In what follows I will examine and criticize Conee's attempt to give a definition of instrumental value that accounts for preventions. I will then propose an analysis of extrinsic value that accounts for instrumental and preventive value. Finally, I will argue that my analysis also accounts for signatory and contributory value.

SOME FLAWED DEFINITIONS

Following Conee, I will take the following statement by Charles Baylis as a starting point: "Anything is instrumentally good in so far as it is a causal factor in the production of something which is good" (Baylis 488). This suggests the following analysis of instrumental value (the "naive view"):

NV: s is instrumentally good $=_{df}$. s is a causal factor in the production of something good.

NV suffers from several defects. The first is the vague term 'good' in its *definiens*. What sort of goodness does Baylis have in mind? Intrinsic goodness? Overall goodness? Perhaps either intrinsic or instrumental goodness?

No matter which we choose, NV is false. If something is a causal factor in the production of something intrinsically good, for example, it may also be a causal factor in the production of something intrinsically very bad. This point applies no matter how we interpret 'good' in NV. To make the point clear: suppose we interpret 'good' in NV as 'overall good.' Suppose that x causes both y and z; y has an overall value of −1,000, and z has an overall value of +10. In such a case, x is a causal factor in the production of z; and z has an overall value of +10, so z is overall good. So according to NV, x is instrumentally good. Yet in this case, x clearly is not instrumentally good, since besides causing z, it causes y, which is overall very bad.

The problem with NV is that it fails to make use of the notion of something's *total causal consequence*.[5] To explain this notion, we should first make some preliminary decisions. Let's assume that the bearers of intrinsic and extrinsic value are states of affairs, or propositions.[6] Assume also that possible worlds are conjunctions of states of affairs, and that for every world w and state of affairs s, either s or not-s is true at w. The total causal consequence of a state of affairs s at w can be defined as the conjunction of all states of affairs at w that are caused by s. Next we need to calculate the intrinsic value of the total causal consequence of s at w. There are different views as to how we should calculate the intrinsic value of a conjunctive state of affairs. Personally, I am partial to the standard view that the intrinsic value of a conjunctive state of affairs is equal to the sum of the basic intrinsic values of its conjuncts.[7] Others, however, might reject this view as too simplistic; perhaps there could be a conjunctive state of affairs whose intrinsic value is greater than the sum of the basic intrinsic values of its conjuncts – an "organic unity."[8] Henceforth I will assume the standard view, for two reasons: (i) it is simple; (ii) I do not know how an advocate of organic unities would calculate the value of a conjunctive state of affairs. The following formula gives us the instrumental value of s:

BIV: For any state of affairs s and world w, the instrumental value of s at w = the intrinsic value of the total causal consequence of s at w.

It should be obvious that the instrumental value of a state of affairs must be relativized to a world, since a single state of affairs will have different effects at different worlds.

BIV solves the problems I noted for NV. I suspect that something like BIV is what lies behind the remarks on instrumental value made by Baylis, Brandt, Frankena, Hospers, and Pojman. But even BIV is not true; for as Conee notes, states of affairs may have *preventive* value. "It is clear that we sometimes take things to have instrumental merit because of what they prevent as well as what they accomplish" (Conee 111). Vaccinations, for example, are considered to be instrumentally good in virtue of preventing the evils that would otherwise be brought about by disease. It seems quite plausible to suppose that vaccinations would be good even if those vaccinated go

on to live lives with zero intrinsic value, for some of those vaccinated would have been much worse off had they not been vaccinated. Thus some states of affairs seem to be instrumentally good without having intrinsically good total causal consequences.

This leads Conee to propose the following definition:

> CIV: s is instrumentally good $=_{df}$. s causally contributes to bringing about something overall good or preventing something overall bad. (Conee 111)

Conee thinks this gives us a concept of interest. But CIV suffers from some of the same defects as NV, since the first disjunct of the *definiens* of CIV is relevantly similar to the *definiens* of NV. A further problem which CIV is that it defines instrumental value in terms of overall value. If we understand something's overall value, as Conee seems to,[9] to be the sum of its intrinsic and instrumental values, then CIV is rather uninformative. We can't understand CIV unless we already understand what instrumental value is. If we iron out these problems, we get something like the following:

> CIV′: For any state of affairs s and world w, s is instrumentally good at w $=_{df}$. s's total causal consequence at w is intrinsically good or s causally contributes to preventing something intrinsically bad from occurring at w.

CIV′ still suffers from serious defects; for contributing to preventing something intrinsically bad does not seem sufficient to make something instrumentally good. Suppose a comet is heading towards earth. The comet will land on one of three places, depending on which way the wind happens to be blowing. If the wind is blowing north, then the comet will land on Amherst, Massachusetts; many people will die, so the comet's crashing will cause states of affairs with intrinsic value of (say) −10,000. If the wind is blowing south, then the comet will land on Springfield, Massachusetts; again, the intrinsic value of the states of affairs caused will be −10,000. If the wind is blowing east, the comet will land harmlessly in the ocean; the intrinsic value of the states of affairs caused will be zero. Suppose the wind blows north and the comet lands on Amherst. It seems reasonable to suppose that the wind's blowing north prevented the comet from landing on Springfield; thus, the wind's blowing north prevents many

intrinsically bad states of affairs. If CIV′ were true, it would be instrumentally good for the wind to blow north. But this is obviously not instrumentally good.

Perhaps we might try to revise CIV′, as follows:

CIV″: For any state of affairs s and world w, s is instrumentally good at w $=_{df}$. s's total causal consequence at w is intrinsically good or s causally contributes to preventing something intrinsically worse than s's total causal consequence at w from occurring at w.

This revision appears to handle the counterexample I gave to CIV′. Given the three possibilities (i) the comet lands on Amherst, (ii) the comet lands on Springfield, and (iii) the comet lands in the ocean, if (i) were to obtain, it would not prevent anything intrinsically worse than the total causal consequence of (i). So according to CIV″, (i) turns out not to have positive instrumental value, in accord with our intuitions.

But CIV″ still suffers from a fatal defect. Suppose that the comet lands on Amherst in 1996, but had it not landed on Amherst, it would have landed harmlessly in the ocean. Furthermore, suppose that had the comet landed in the ocean, the citizens of Amherst would have gone on to live prosperous and happy lives for many years. The town would have thrived and grown for centuries until 2496 when, tragically, another comet would have come along and destroyed it. As it would have happened, due to a few additional deaths caused by the growth of the town, the intrinsic value of the states of affairs caused by the comet of 2496 would have been −11,000. According to CIV″, since the comet's landing on Amherst in 1996 had a total causal consequence with intrinsic value of −10,000 but (it seems reasonable to suppose) prevents a state of affairs with intrinsic value of −11,000, it would be instrumentally good for the comet to hit Amherst in 1996. As I've described the situation, this would clearly not be instrumentally good; it prevents many people from living very happy lives.

I conclude that CIV″ is a failure. There seems to be something of interest lurking around here, though. "Intuitively, the idea is that it is equally creditable to move the world upward some amount on the overall value scale from what would have been, whether or not the

change introduces something better than neutral" (Conee 116). We might try to capture this idea in a definition like the following: 's is instrumentally good $=_{df}$. s's total causal consequence is intrinsically better than what s overall prevents.'[10] Such a definition would account for the problem I've noted for CIV″. But I don't want to stop here. I think that Conee's intuitive idea can be fleshed out into a definition not only of instrumental value, but of extrinsic value generally. In the following section I will motivate, present, and defend a "contextualist" account of extrinsic value and show how it implies the definition of instrumental value just presented.

A CONTEXTUALIST ACCOUNT OF EXTRINSIC VALUE

When we calculate the intrinsic value of a world, or a life, we need to look only at the intrinsic values of states of affairs that are true in that world, or involving that life. We don't need to know what happens at other worlds. But in order to have the concept of extrinsic value capture Conee's intuitive idea, the extrinsic value of a state of affairs at a world will depend not only on the intrinsic values of states of affairs true at that world, but at other relevant worlds as well. For we want to know if what the state of affairs leads to is overall better than what *would have* happened if it hadn't obtained. Without considerations of other possible worlds, it's impossible to factor in the extrinsic values of preventions. So we need to know which other world or worlds to consider when calculating something's extrinsic value.

Unfortunately, there's no mechanical way to determine what world to consider, because what would have happened if some state of affairs hadn't obtained depends on what factors we keep fixed from world to world. For example: a grenade falls in the middle of a group of people. Person x heroically jumps on it, sacrificing himself to save the group. There's a clear sense in which his jumping on it was an extrinsically good thing because of the suffering it prevented – even if the saved lives end up having no intrinsic value from then on. But maybe there was another person in the group – call her y – who was willing to jump on the grenade, but waited to see if x would do it first. If x hadn't jumped on the grenade, y would have. The intrinsic value of the world had y jumped on the grenade would

have been the same as if x had jumped on it. Then x's jumping on the grenade fails to move the world up on the overall value scale, and so according to Conee's intuitive notion, it's not even extrinsically good. Is x's jumping on the grenade extrinsically good or not? Is there any *absolute* answer to this question?

I think that there is not; rather, the answer will depend upon what world we suppose would have obtained had x not jumped on the grenade; and this, in turn, depends (at least in part) upon the context of utterance. When we say that x's jumping on the grenade was extrinsically good, we're supposing that the following counterfactual statement is true: 'if x hadn't jumped on the grenade, many would have died.' Thus we must be comparing its effects to the effects in a world where everyone else acts just as they do in the actual world – so nobody jumps on the grenade. It's extrinsically good compared to that world; it makes the world better than that world. When we say that x's jumping on the grenade wasn't extrinsically good, we're supposing that the following counterfactual statement is true: 'if x hadn't jumped on the grenade, y would have.' Thus we are comparing its effects to the ones at the world where y jumped on the grenade – a world no worse than the actual world. Given appropriate contexts of utterance, both answers may be right.

These considerations demonstrate an obvious point – that what would have happened had some state of affairs not obtained is not an absolute matter. Contextual factors enter into determining what would have happened; they determined which world (or worlds) is 'closest' to the actual world. If what would have happened is dependent on the context, then so is the extrinsic value of a state of affairs; it will vary not only from world to world, but also from context to context. So the extrinsic value function will be a function, not simply from pairs of states of affairs and worlds to numbers, but from triples of states of affairs, worlds and *contexts* to numbers (or alternatively: from states of affairs to triples of worlds, contexts, and numbers). Defining extrinsic value in this way, taking contextual factors into account, seems to me to be a promising approach. Appealing to a shift of context would enable us to explain away apparently troublesome "disagreements" (such as one that could arise in the grenade example above) over whether something is extrinsically good or not.

Here is how I think extrinsic value should be calculated. Suppose s is a state of affairs that occurs in world w. Let W be a function that determines what world would have obtained, given a context c, had s not occurred – i.e., the closest world to w, given c, where s doesn't occur; we can represent this as $W\langle\sim s,w,c\rangle$. To determine the extrinsic value of s at w given c, or $EV\langle s,w,c\rangle$, first we determine all the states of affairs *other than s* that obtain in w but not in $W\langle\sim s,w,c\rangle$ – call their conjunction 'the total consequence of $\langle s,w,c\rangle$' – and calculate the intrinsic value of the total consequence of $\langle s,w,c\rangle$. (Note that the notion of 'total consequence' is not equivalent to the notion of 'total causal consequence' – see note 5.) The intrinsic value of the total consequence of $\langle s,w,c\rangle$ is equal to the sum of the basic intrinsic values of its conjuncts. Then we determine the conjunction of all the states of affairs that obtain in $W\langle\sim s,w,c\rangle$ but not in w – call this 'what $\langle s,w,c\rangle$ prevents' – and let the intrinsic value of what $\langle s,w,c\rangle$ prevents be the sum of the basic intrinsic values of its conjuncts.[11] Here, then, is the formula that determines the extrinsic value of $\langle s,w,c\rangle$:

BEV: For any state of affairs s, world w, and context c, $EV\langle s,w,c\rangle$ = IV(the total consequence of $\langle s,w,c\rangle$ – IV(what $\langle s,w,c\rangle$ prevents).[12]

I think BEV comes close to capturing Conee's intuition (though it is formulated in terms of intrinsic value, not overall value).[13] It seems to yield the right results in the grenade example above. Let's let X = the state of affairs consisting in x's jumping on the grenade, and Y = the state of affairs consisting in y's jumping on the grenade; X obtains at @, but Y doesn't. Let C1 = the context in which the counterfactual 'if X hadn't obtained, Y would have obtained' is true. Let C2 = the context in which the counterfactual 'if X hadn't obtained, many would have died' is true. Suppose that those whose lives were saved go on to have lives with zero intrinsic value; that their deaths would have had intrinsic value of −200; that x's death has intrinsic value of −50, and that y's death would also have had intrinsic value of −50. If we apply BEV to this situation, we find that $EV\langle X,@,C1\rangle$ = zero; for IV(the total consequence of $\langle X,@,C1\rangle$ = IV(what $\langle X,@,C1\rangle$ prevents) = −50. We also find that $EV\langle X,@,C2\rangle$ = +150; for IV(the total consequence of $\langle X,@,C2\rangle$) = −50, while

IV (what ⟨X,@,C2⟩ prevents) = −200. Thus, given context C2, X is extrinsically good, but given context C1, it isn't. That's just the result we wanted.

SIGNATORY AND CONTRIBUTORY VALUE

A further virtue of BEV is that it accounts for signatory and contributory value. Signatory value can be explained by the following fictional story: suppose Shaquille is a basketball player who injures his knee on the court. He goes to the doctor to find out if he has torn ligaments. If his knee has torn ligaments, the world will be a worse place. He won't be able to enjoy playing basketball for several weeks; he will have to undergo a painful operation and difficult rehabilitation; he will will make only fifteen million dollars the next season, not twenty million; and so on. The doctor x-rays his knee. As it happens, the x-ray comes back, with no sign of torn ligaments. We might say that it's a good thing that the x-ray looks the way it does. But of course it's not intrinsically good that it looks that way; nor is it instrumentally good. Rather, its value comes from what it *signifies* – viz., that Shaquille doesn't have torn ligaments.

BEV accounts for this quite nicely. Let X = the state of affairs consisting in Shaquille's x-ray coming out with no sign of torn ligaments. Suppose X occurs in @. In a context in which we would say that X is good, we would be assuming a number of things: most importantly, that if X had not occurred, Shaquille would have had torn ligaments. Call this context C. So W⟨∼X,@,C⟩ is a world where Shaquille has torn ligaments. To determine whether X is extrinsically good, we first determine the sum of the basic intrinsic values of all the states of affairs that are true at @ but not at W⟨∼X,@,C⟩. Assuming (implausibly) that Shaquille lives a totally mediocre life, this would be zero. Then we determine the sum of the basic intrinsic values of all the states of affairs that are true at W⟨∼X,@,C⟩ but not at @. Some of these will be states of affairs where Shaquille experiences great pain, so it seems likely that the sum will be a negative number – say −100. Given BEV, the extrinsic value of ⟨X,@,C⟩ is +100; thus it is a very good thing that the x-ray comes out negative. The fact that BEV does not make the extrinsic value of a state of affairs depend specifically on what it *causes* allows BEV to account for a

wider variety of kinds of extrinsic value than definitions that involve causal relations.

One might be skeptical of the claim that BEV accounts for signatory value. It might be objected that the claim that the x-ray's coming out negative is a good thing does not imply the claim that things would have been worse had the x-ray been positive.[14] We might truly claim that it was a good thing the x-ray was negative, the objection goes, but later find out (as a result of further tests) that Shaquille's leg has problems that the x-ray failed to detect. If it could be true that the x-ray has positive signatory value, yet the world is no better than it would have been otherwise, then this would indeed be a serious objection to my contention that BEV accounts for signatory value.

I don't find this objection very troubling. It seems to me that in order for a sign to have positive signatory value, what it signifies must be *true*; and I think that careful reflection should lead the reader to the same conclusion. If I were to Shaquille's position in the above story, I would find no inclination to call the initial x-ray result good unless what it signified – that there was no serious injury to the leg – were true. Upon hearing the grim results of the later tests, I would be inclined to retract any earlier remarks I might have made about the positive value of the x-ray results. Imagine the doctor coming into the room and saying to Shaquille, "I have some good news and some bad news. The good news is, the x-ray shows no sign of injury! The bad news is, the MRI shows that you'll be on crutches for three months." Shaquille could only take the doctor's remarks as a joke. There's obviously *no* good news here; the x-ray results are clearly worthless.

Here is one more example in support of my intuition. There's an old adage that goes something like "red sky at night, sailor's delight; red sky in morning, sailors take warning." Presumably a red sky in the morning usually indicates that there will be bad weather that day; let's just suppose that is true. Suppose two sailors wake up to find the sky bright red. One says to the other: "Look at how red the sky is! That's a bad sign." Suppose, however, that the weather turns out to be perfect for sailing. I think that the sailors would agree that the red sky had not been a bad thing after all. Perhaps there is another kind of signatory value for which BEV cannot account; but I think it at least accounts for one interesting kind.

Finally we have contributory value. C. I. Lewis defines contributory value as follows: "Let us call the value assignable to any transitory experience not – or not merely – by reason of the quality it immediately presents out on account of its contribution to some larger whole of experience, or to life altogether, the *contributory* value of it" (Lewis 68). Ignoring Lewis' talk of "experiences," here is what I think contributory value amounts to: we might adopt an axiology according to which some of the states of affairs with basic intrinsic value are complex states of affairs with other states of affairs as constituents, and where either (i) the constituent states of affairs are intrinsically worthless or (ii) the intrinsic value of the whole is greater than the sum of the basic intrinsic values of its constitutions. In principle, there seems to be no reason why we could not adopt such an axiology. (Perhaps someone might adopt an axiology according to which a life, or the conjunction of all the states of affairs about the life of a person, has basic intrinsic value.) The constituents themselves might have no intrinsic value; yet we might wish to say they have some sort of value, since if the constituents of a complex state of affairs do not obtain, the whole does not obtain either. The value a constituent has in virtue of contributing to the whole is its contributory value.

I think BEV has no problem accounting for contributory value, if there is such a thing.[15] Let R and S be states of affairs with no intrinsic value that obtain at world @; let R&S be a state of affairs with basic intrinsic value of +10. (We might call R&S an "organic unity.") Suppose that neither R nor S is a constituent of any other state of affairs with basic intrinsic value at @, and that R neither causes nor is a sign of any state of affairs with basic intrinsic value at @. To determine the extrinsic value of R, we must determine what would have happened had R not obtained. Suppose the context C implies that had R, and all the states of affairs of which R is a constituent, not obtained, they wouldn't have been "replaced" by any intrinsically valuable states of affairs; then the intrinsic value of what $\langle$R,@,C$\rangle$ prevents will be zero. The total consequence of $\langle$R,@,C$\rangle$ will include only states of affairs (other than R) with R as a constituent; the only one of these with basic intrinsic value is R&S, so the intrinsic value of the total consequence of $\langle$R,@,C$\rangle$ is

+10. Thus BEV implies that the extrinsic value of ⟨R,@,C⟩ is +10; it seems reasonable to call this its contributory value.

CONCLUDING REMARKS

I have discussed a few of the most famous alleged types of extrinsic value. Of course, there are countless sorts of value that I haven't discussed here, such as "inherent" value[16] and "product" value.[17] I cannot possibly hope to discuss every interesting kind of value. It might be that there is a sort of (moral) extrinsic value that presents a problem for BEV. I am content to make the modest claim that BEV accounts for several interesting types of extrinsic value: instrumental, preventive, signatory, and contributory.

I haven't specifically defined the expressions 'x has instrumental value,' 'x has signatory value,' and so forth. Perhaps it is possible to give distinct definitions of each of these expressions, but I don't think it is helpful or necessary to do so. When we are interested in, say, the signatory value of s at w, what happens is that the context c gives us a value for W⟨∼s,w,c⟩ in which all the states of affairs that are true at w but not W⟨∼s,w,c⟩ are ones of which s is a sign. If we wanted a formula, we could give this:

SV: For any state of affairs s, world w, and context c, the signatory value of ⟨s,w,c⟩ = [IV(what ⟨s,w,c⟩ is a sign of) – IV(s)] – IV(what ⟨s,w,c⟩ prevents).

But this is really not to give anything new; for in any instance where we are asking about signatory value, the total consequence of ⟨s,w,c⟩ would be identical to what ⟨s,w,c⟩ is a sign of.

It is worthwhile to notice that the way of calculating extrinsic value described here implies that a world may contain states of affairs that are extrinsically good while no states of affairs at that world are intrinsically good. This may seem counterintuitive; many have thought that in order for something to be extrinically good, it must bear some relation to something with positive intrinsic value.[18] That is, one might think that extrinsic value is calculated in the following way (reminiscent of BIV):

BEV2: For any state of affairs s, world w, and context c, EV⟨s,w,c⟩ = IV(total consequence of ⟨s,w,c⟩).

This certainly gets at a different conception of extrinsic value than BEV. Perhaps there are, in fact, two useful conceptions of extrinsic value – one that takes preventive value into account, and one that doesn't. But I find this unlikely. The intuition that something's being extrinsically good implies the existence of something with positive intrinsic value can, I think, be accounted for by BEV. The idea is that when we require a state of affairs to lead to (or to be a sign or a part of) something with positive intrinsic value in order to be extrinsically good, we are simply insisting upon a context in which what the state of affairs prevents has intrinsic value of zero. If what the state of affairs prevents has zero intrinsic value, then in order for the state of affairs to have positive extrinsic value according to BEV, its total consequence must have positive intrinsic value – so something intrinsically good must exist at that world.

At any rate, since BEV defines extrinsic value in terms of intrinsic value, it does account for an intuition similar to the one just described – the intuition (which I share) that unless there is such a concept as intrinsic value, talk of extrinsic value makes no sense. To put it another way, unless something *at some world or other* has intrinsic value or disvalue, nothing at any world can have extrinsic value or disvalue.[19]

APPENDIX: A BASIC PROBLEM FOR BEV

There's a minor problem with BEV involving basic intrinsic values. The following example should make the glitch apparent. Suppose that hedonism is true, and that the only states of affairs with basic intrinsic value consist in a particular person's (*de re*) experiencing a definite amount of pleasure or pain at a time. Suppose that S is a state of affairs with positive basic intrinsic value that obtains at world @ – Jones' experiencing 12 units of pleasure at noon. Suppose that S's truth at @ implies the truth at @ of another state of affairs T that has intrinsic value, but not *basic* intrinsic value: *someone's* experiencing 12 units of pleasure at noon. We might say that S "bestows" T with whatever intrinsic value it has – viz., +12. To determine the extrinsic value of T according to BEV, we would have to include S as part of its total consequence; for T is distinct from S, and S is true at the world w where T occurs, but not at W⟨∼T,w,c⟩ (for any

C). Since S has basic intrinsic value, it would boost the intrinsic value of the total consequence of T, and hence boost the extrinsic value of T. This seems wrong; we don't want S to be factored into the calculation of the extrinsic value of T.

To make this a bit more clear, suppose we add the following details to the above example: T causes R, a state of affairs which has a basic intrinsic value of +10; T causes no other state of affairs with intrinsic value; and the context, C, implies that had T not occurred, nothing with any intrinsic value would have occurred – i.e., IV(what ⟨T,@,C⟩ prevents) = zero. In such a situation, we'd like to say that T's extrinsic value is +10. But according to BEV, the total consequence of T includes both R and S; so we add the basic intrinsic values of R and S to get the intrinsic value of the total consequence of ⟨T,@,C⟩. Since the basic intrinsic values of R and S add to +22, and since what T prevents has zero intrinsic value, BEV entails that T's extrinsic value would be +22.

The problem is easy to fix. We revise the notion of the total consequence of ⟨s,w,c⟩ to include s; so the total consequence of ⟨s,w,c⟩ will be the conjunction of *all* states of affairs true at w that are not true at W⟨∼s,w,c⟩. Then we revise the formula for calculating extrinsic value as follows:

BEV3: For any state of affairs s, world w, and context c, EV⟨s,w,c⟩ = [IV(the total consequence of ⟨s,w,c⟩ – IV(s)] – IV(what ⟨s,w,c⟩ prevents).

I think this solves the problem. In the above example that posed a problem for BEV, the intrinsic value of S will be factored into the intrinsic value of the total consequence of T, but it will be factored out when we subtract the intrinsic value of T. The extrinsic value of T will be +10, just as we'd hoped.

An astute reader may have realized that BEV3 is in fact equivalent to the following slightly more elegant formula:

BEV3′: For any state of affairs s, world w, and context c, EV⟨s,w,c⟩ = [IV(w) – IV(s)] – IV(W⟨∼s,w,c⟩).

While BEV3′ is more elegant, it is perhaps more difficult to see why it yields the right results. But indeed, it does, since all and only the states of affairs that are true at w but not part of the total consequence of ⟨s,w,c⟩ are states of affairs that are true at W⟨∼s,w,c⟩ but not part of what ⟨s,w,c⟩ prevents. Hence their (basic) intrinsic values are included in both IV(w) and IV(W⟨∼s,w,c⟩), and thus do not affect the calculation.

NOTES

[1] That intrinsic and instrumental goods are the only kinds of goods seems also to be implied by remarks on intrinsic value by Judith Thomson (Thomson, p. 131).

[2] Frankena does not, however, claim that intrinsic and instrumental goods are the only kinds of goods.

[3] Fred Feldman mentions this possibility (1986, p. 26): "We sometimes hear a doctor say, 'that's good', when looking at an X-ray photograph. Perhaps he sees something that he takes to be a sign of good health."

[4] Robert Olson seems to hold a similar view about extrinsic, but not necessarily instrumental, value: "a nonintrinsic good is something valuable by virtue of its relationship to an intrinsic good" (Olson, p. 367). See also the Frankena quotation (above, p. 2) and the Charles Baylis quotation (p. 4).

[5] I make a distinction between something's total consequence and something's total *causal* consequence. Something's total causal consequence includes any state of affairs caused by it; something's total consequence may include other states of affairs as well. The notion of a total causal consequence is relevant when defining instrumental value, while to define extrinsic value we need the more general notion of total consequence; this notion is introduced below.

[6] As I understand states of affairs, there are many different sorts. There are some that obtain at some times but not others; John's eating a hamburger is one such state of affairs. There are others, such as John's eating a hamburger at 12 noon on September 10, 1996, whose truth values never change over time. I will be more concerned with the extrinsic values of the latter type of state of affairs, simply because I think it is these states of affairs to which we most often attribute value.

It has been objected, by an anonymous referee for this journal, that "it is not a state of affairs itself that has value, but rather that its obtaining does. (The state of affairs of everyone being extremely happy exists at this world, but that doesn't make this world a tremendously good one; for it doesn't obtain here.)" Indeed, it is true that a state of affairs does not affect the intrinsic value of a world in which it is false; however, this is an objection to states of affairs' being the bearers of value only if we suppose that in order to find the intrinsic value of a world, we simply add the basic intrinsic values of all the states of affairs that *exist* there. This seems implausible; a more plausible way to determine the intrinsic value of a world would be to add the basic intrinsic values of the states of affairs that are *true* at that world. Besides, we might just as well wonder what an "obtaining" is. Is it a state of affairs itself? If so, then the objection clearly carries no weight.

[7] The notion of "basic" intrinsic value is necessary in axiology to avoid double-counting. On the need for basics, see Harman (1967).

[8] The existence of organic unities was alleged by Moore (1903); the issue is of course very deep and puzzling, and I don't wish to discuss it here.

[9] Conee claims that "the right concept to use is what we call being 'good on the whole' or 'overall good.' . . . In the case of that notion, plainly it 'does not matter' whether the impact is positive because of intrinsic or instrumental factors" (Conee, p. 105).

[10] Here I am indebted to an anonymous referee for this journal, though it should be noted that this definition is not quite identical to the definition suggested by the referee.

[11] Note that the sense of 'prevents' utilized here is a very broad notion – probably

much broader than the ordinary use of the word. In the narrower, ordinary sense, to prevent something is to cause it not to occur; but one state of affairs may prevent another from occurring in the broader sense introduced here without *causing* it not to occur.

[12] BEV is similar to a definition of extrinsic value given by Fred Feldman: "The extrinsic value for S of P = the difference between the intrinsic value for S of the life S would lead if P is true and the intrinsic value for S of the life S would lead if P is false" (1992, p. 150). There are some differences, however. Feldman defines extrinsic value for lives, and he does not mention contexts. Most importantly, his definition is false. Suppose John experiences some pain that has no bad consequences. His life is made worse by that pain, so Feldman's definition entails that the pain is extrinsically bad for John.

[13] There is a technical glitch with BEV involving basic intrinsic values. For the sake of continuity, I confine the problem and its solution to the appendix below.

[14] This objection was raised by an anonymous referee for this journal.

[15] I think that if there is such a thing as contributory value, there must be organic unities. If this is correct, then the summative principle (see p. 6) I have been assuming throughout must be rejected. (Here I am indebted to an anonymous referee for this journal.)

[16] See C. I. Lewis (1955, p. 69).

[17] See Thomson (1990, p. 132n).

[18] For example, Lafleur writes: "if anything is properly described as an extrinsic good, there must be intrinsic goods which determine its nature" (Lafleur, p. 211). Brandt agrees: "It would seem, however, that if anything is desirable at all, it must be because some facts or events . . . are intrinsically desirable" (Brandt, p. 302).

[19] I am indebted to Fred Feldman and an anonymous referee for this journal for helpful comments on earlier versions of this paper.

REFERENCES

Baylis, C. (1958): 'Grading, Values, and Choice', *Mind* 67, 485–501.

Bentham, J. (1948): *An Introduction to the Principles of Morals and Legislation*, New York: Hafner Publishing Co.

Brandt, R. (1959): *Ethical Theory: The Problems of Normative and Critical Ethics*, Englewood Cliffs, NJ: Prentice-Hall Inc.

Conee, E. (1980): 'Pleasure and Intrinsic Goodness', Ph.D. Dissertation, University of Massachusetts at Amherst.

Feldman, F. (1986): *Doing the Best We Can: An Essay in Informal Deontic Logic*, Dordrecht, Holland: D. Reidel Publishing Co.

Feldman, F. (1992): *Confrontations with the Reaper: A Philosophical Study of the Nature and Value of Death*, New York: Oxford University Press.

Frankena, W. (1963): *Ethics*, Englewood Cliffs, NJ: Prentice-Hall, Inc.

Harman, G. (1967): 'Toward a Theory of Intrinsic Value', *The Journal of Philosophy* 64, 792–804.

Hospers, J. (1961): *Human Conduct*, New York: Harcourt, Brace, & World.

Lafleur, L. (1954): 'The Meanings of Good', *Philosophy and Phenomenological Research* 15, 210–221.

Lewis, C.I. (1955): *The Ground and Nature of the Right*, New York: Columbia University Press.
Mill, J.S. (1957): *Utilitarianism*, New York: Macmillan Publishing Co.
Moore, G.E. (1903): *Principia Ethica*, New York: Cambridge University Press.
Olson, R. (1967): 'The Good', in P. Edwards (ed.), *The Encyclopedia of Philosophy*, vol. III, New York: MacMillan Co. & the Free Press.
Plato. *Republic*, trans. A. Bloom, Basic Books, 1991.
Pojman, L. (1993): *Moral Philosophy: A Reader*, Indianapolis: Hackett Publishing Co.
Thomson, J. (1990): *The Realm of Rights*, Cambridge: Harvard University Press.

Department of Philosophy
University of Massachusetts
Amherst, MA 01002
USA