Philosophical Studies (2006) 128: 441–477
DOI 10.1007/s11098-004-7810-5
© Springer 2006

JOSEPH MENDOLA

INTUITIVE HEDONISM

ABSTRACT. The hoary philosophical tradition of hedonism – the view that pleasure is the basic ethical or normative value – suggests that it is at least reasonably and roughly intuitive. But philosophers no longer treat hedonism that way. For the most part, they think that they know it to be obviously false on intuitive grounds, much more obviously false on such grounds than familiar competitors. I argue that this consensus is wrong. I defend the intuitive cogency of hedonism relative to the dominant desire-based and objectivist conceptions of well-being and the good. I argue that hedonism is still a contender, and indeed that our current understanding of commonsense intuition on balance supports it.

The hoary philosophical tradition of hedonism – the view that pleasure is the basic ethical or normative value – suggests that it is at least reasonably and roughly intuitive. But philosophers no longer treat hedonism that way. For the most part, they think that they know it to be obviously false on intuitive grounds, much more obviously false on such grounds than familiar competitors. I will argue that this consensus is wrong.[1] I will defend the intuitive cogency of hedonism relative to the dominant desire-based and objectivist conceptions of well-being and the good. I will argue that hedonism is still a contender, and indeed that our current understanding of commonsense intuition on balance supports it.

The success of such an argument may seem quite unlikely. For instance, here is Sumner's relatively sympathetic evaluation of hedonism's prospects:

> Time and philosophical fashion have not been kind to hedonism. Although hedonistic theories of various sorts flourished for three centuries or so in their congenial empiricist habit, they have all but disappeared from the scene. Do they now merit even passing attention, for other than nostalgic purposes? Like endangered species, discredited ideas do sometimes manage to make a comeback (Sumner, 1996).

Of course, always at issue is whose intuitions count. The fact that hedonism was a serious contender within philosophy for much longer than even the three centuries Sumner notes suggests that it might still have a life outside of the trained intuitions of contemporary analytic moral philosophers. Nevertheless, it is those intuitions I will address. Some of my strategy will be to delimit the kind of hedonism which I will defend, which will soften its conflict with our trained intuition. But most of my strategy will be to undercut and adjust the intuitions which seem to discredit hedonism, by appeal to other intuitions which are more favorable and forceful.

Section 1 distinguishes the form of hedonism which I will defend from others in its near vicinity. Sections 2, 3, and 4 argue that hedonism is appropriately intuitive, despite recent fashion.

1

The hedonism I will defend is at once very crude, which may make it seem especially hard to defend on intuitive grounds, and yet deployed in a specific way which somewhat eases that task. Let's begin with the crudity: Focus on the pleasantness of physical pleasures and the painfulness of physical pains. I believe that the pleasantness of physical pleasure *is* a kind of hedonic value, a single homogenous sensory property, differing merely in intensity as well as in extent and duration, which is yet a kind of goodness.[2] Likewise for pain and hedonic disvalue. Alternatively and less idiosyncratically, one might claim that the sensory pleasantness of physical pains has a value which would be in that sense hedonic. It would not matter for our purposes. While, physical pains and crude physical pleasures provide the paradigm examples of this hedonic value and disvalue, other "higher" pleasures and pains may involve attenuated experience of the same kind.

This is reminiscent of the classical hedonism of Bentham, which is the sort now most widely ridiculed. The views of the few contemporary hedonists deploy more sophisticated con-

ceptions of pleasure or enjoyment (Sidgwick, 1907; Edwards, 1979),[3] and even these more sophisticated forms are widely thought untenable. At least since Sidgwick, it has been perhaps the dominant view among even hedonists that there is not a single homogeneous quality of sensation characteristic of all pleasures or of all pains. I believe that some real phenomenal differences between pains and pleasures can be delivered by other felt elements than hedonic value, and I do not insist that hedonic value or disvalue provide a phenomenal reduction of or necessary condition for all pains and pleasures. Still there is no questioning the relatively simple crudity of the basic phenomenal properties which I will defend here as being the basic normative properties, or at least as being the natural ground of the basic normative properties. Contemporary defenders of hedonism usually adopt what Sumner has called an externalist conception of pain and pleasure, in which pain is characterized as involving a certain sort of characteristic aversion to a wide range of phenomenal experiences.[4] But hedonic value is internalist in this sense and phenomenally homogeneous.

Still, even my crude hedonism does not identify pleasure itself with positive value, but rather merely considers the positive hedonic tone characteristic of physical pleasures to be the sole irreducible kind of positive value. And internalist conceptions of pleasure and pain do have certain intuitive advantages, well-summarized by Sumner:

> The attitude model [of pain, which identifies it with all aversive phenomenal states,] can make no sense at all of the testimony of lobotomized subjects who say that they continue to feel pain but are no longer averse to it. More seriously, it also runs foul of the perfectly obvious fact that pain is not the only physical feeling to which we are (normally) averse. Think for a moment of the many physical symptoms which, when persistent, can make our lives miserable: nausea, hiccups, sneezing, dizziness, disorientation, loss of balance, itching, ''pins and needles'', ''restless legs'', tics, twitching, fatigue, difficulty in breathing, and so on. While none of these is quite the same as physical pain, we experience each as intrinsically disagreeable. The attitude model simply obliterates these categorial boundaries by treating all these states indifferently as pain.[5]

But, on the other hand, my crude conception has exactly corresponding disadvantages. It deploys phenomenal hedonic value and disvalue of a simple sort even to characterize the phenomenal value component of itching or pins and needles. And the fact that we sometimes do not mind hedonic value and sometimes do mind other phenomenological states is not obviously an intuitive advantage for my account. Let me put it this way: Perhaps it is reasonable to follow Casell and distinguish between pleasure and pain on the one hand and enjoyment and suffering on the other, where pain and pleasure are mere sensations but enjoyment and suffering involve, perhaps only, the attitudes which externalism customarily deploys (Casell, 1983). But then pain which is not suffering and pleasure which is not enjoyment may seem intuitively irrelevant to our well-being or good. And hence a plausible hedonism may seem of necessity to be a hedonism of enjoyment and suffering, and not of pleasure and pain, let alone of the homogeneous and crude hedonic tone which I deploy.

Nevertheless, I will defend my crude hedonism as sufficiently and appropriately intuitive. One thing which helps is the relative specificity of the role in which I would defend it. First, this hedonism is in no interesting sense psychological hedonism. I make no claim that hedonic value is our ultimate goal or desire, and indeed no claim that it plays any specific motivational or psychological role at all. And even within normative hedonism, we must heed some distinctions. First, I deploy hedonic value and disvalue, and not pain and pleasure or suffering and enjoyment, *strictu sensu*. So my view is *hedonism* in though a recognizable still an extended sense. Second, there are a variety of different normative roles in which pleasure or hedonic value may be deployed. It might be thought to be the only thing which is good, or the only thing which is intrinsically or non-instrumentally good. Or it might be thought that one's own pleasure is all that is ultimately *good for one* or all that can constitute one's well-being. Indeed, there may be more than one notion of well-being to which pleasure contributes in different ways, since, as a number of authors have noted, different notions of individual well-being may be differentially appropriate

for different normative purposes, say for the distribution of certain government services or the retrospective evaluation of one's own life (Scanlon, 1998). But in fact I will not deploy hedonic value in any of these exact normative roles.

I will defend hedonic value and disvalue only as providing what one might call *basic* value, which must be augmented, for instance, by a plausible consequentialist account of other sorts of value and obligation and even by distribution-sensitivity to provide a proper ordering of whole states of affairs.[6] But still this view is recognizably within the classical utilitarian tradition. There is a recognizable sense in which I claim that the overall normative status of things rests on what we might call the basic well-being and good of sentient beings, and not their overall or final well-being or good. And there is also a recognizable sense in which I claim that this basic well-being and good is constituted by their pleasure and pain, or more precisely by the crude hedonic value which is a component of their paradigmatic pains and pleasures.

So, despite the complications I have noted, the best place to engage the hedonism I will defend and contemporary intuition is in regard to current discussions of well-being and the good. The following sections probe successively into what I take to be sequentially deeper strata of intuitive objections to normative hedonism. Sections 2 and 3 introduce alternative accounts of well-being, and prominent intuitive concrete cases which seem to undercut hedonism and support the alternatives. Section 4, by reference to W.D. Ross' discussion in *The right and the good* (Ross, 1930), addresses in a general way the intuitions which lie behind these specific cases.

2

Parfit provides the following typography of theories of what is good for one: "On the *Hedonistic Theory*, what would be best for someone is what would give him most happiness… On the *Desire-Fulfilment Theory*, what would be best for someone is what would best fulfill his desires throughout his life… On the *Objective List Theory*, certain things are good or bad for us,

even if we would not want to have the good things or avoid the bad things (Parfit, 1984).'' Griffin adopts a similar organization of extant competitor accounts of well-being in his well-known treatment of that issue (Griffin, 1986).[7] To save words and letters, I will speak simply of ''hedonism'', ''desire-based theories'', and ''objectivism''.

It is important to notice that there is more than one possible form of each type of account. Since hedonism is so widely disparaged, my primary concern throughout this paper will be to defend the intuitive cogency of my crude hedonism relative to dominant desire-based and objectivist accounts of well-being and the good. Also, I cannot hope in the short space of a paper to conclusively refute all alternatives to hedonism. My smaller aim is rather to defend the overall relative intuitive plausibility of hedonism. I hope to definitively establish that hedonism is at least an intuitive competitor, that it cannot properly be summarily dismissed as radically counterintuitive. It is not trendy, but it deserves respect. I also more ambitiously hope to convince you that on balance our current understanding of commonsense intuition suggests that hedonism is the overall most intuitive extant account.

As I have noted, this is not a popular view. Contemporary arguments against hedonism and for alternative conceptions of well-being and the good rest heavily on certain concrete cases, which are deployed to engage our commonsense intuition against hedonism or for its competitors. They have seemed quite convincing.

Concrete cases can be vivid and forceful, and they are a typical methodology throughout contemporary analytic philosophy. Nevertheless, they should be handled with some care. First, it is hard in a very concrete case to avoid elements which distract intuition from the crucial features which it is supposed to test. The very concreteness of cases can make them impure. Second, overall judgments about normative and other philosophical matters may be complex, and require that we balance off competing intuitive inclinations in a complex way. So it is important to consider a suitably wide range of different concrete cases which may pull in different directions, if we are to come to

a plausible overall understanding of our intuitive judgments. We must be wary of hasty over-generalization from single cases.

Unfortunately, perhaps just because hedonism no longer seems like a serious contender, the discussions of hedonism by even very able, conscientious, and relatively sympathetic authors are now less than ideal in these ways. I will begin with the standard cases which seem to undercut hedonism. I will focus our discussion on the negative treatment of views like my own by Griffin and by Sumner, who propose instead recognizably desire-based accounts (see also Carson, 2000). Part of my point will be that their accounts of well-being are unusually able and conscientious, relatively sympathetic to hedonism, and standard. They are a best case. The likely explanation for the problems which we will uncover in their discussions is that crude hedonism like my own no longer seems worth real attention. That is also part of my point.

Sumner says of hedonism that the "two strongest objections to an account of this sort have been nicely summarized by James Griffin."[8] But it is striking that Griffin is effectively done with hedonism by the fourth page of his first chapter.[9] His two objections are two sorts of concrete counterexample. The first objection is deployed against simple, homogeneous hedonic tone accounts such as mine and Bentham's, while the second objection is deployed against more complex hedonisms which identify well-being with "states of mind" of a broader sort. Since my hedonism is so crude, and the second objection would count against crude as well as more complex hedonisms, we can safely ignore that detail.

First, there is Freud in pain. Here is the entire relevant passage from Griffin:

> At the very end of his life, Freud, ill and in pain, refused drugs except aspirin. "I prefer", he said, "to think in torment than not to be able to think clearly". But can we find a single feeling or mental state present in both of Freud's options in virtue of which he ranked them as he did? The truth seems, rather, that often we just rank options, period. Some preferences – Freud's seems to be one – are basic. That is, preferences do not always rest upon our judgments about the quantity of some homogeneous mental state found in, or produced by, each option.[10]

First point: A normative hedonist need not deny that we prefer other things than pleasure. I certainly do not deny that we have non-hedonic preferences, even if we ignore complexities like the instrumentality of some desires. I just claim that such preferences can be mistaken. And notice that this paragraph mostly just reminds us that Freud had a non-hedonic preference. Griffin presents this as an objection to a certain conception of utility in particular, which might perhaps plausibly be defined by reference to individual preference. But the problem is that this paragraph constitutes Griffin's entire attack on a simple hedonist notion of well-being, with which utility has been linked on his preceding page.

And there is a second point: While Griffin is a desire-based theorist, he is an *informed* desire-based theorist. He thinks that desires must be corrected by full information to have proper normative weight. Two pages after the paragraph I have quoted, he rejects the "actual-desire" account in this way: "[N]otoriously, we mistake our own interests. It is depressingly common that even when some of our strongest and most central desires are fulfilled, we are no better, even worse, off. Since the notion we are after is the ordinary notion of "well-being", what must matter for utility will have to be, not persons' actual desires, but their desires in some way improved. The objection to the actual-desire account is overwhelming."[11] But if the actual-desire account is false, then a simple appeal to contrary preference cannot establish the falsity of hedonism.

Perhaps my reading of Griffin's case is unsympathetic. Perhaps his central thought was not that Freud simply had these preferences, but that he "ranked" the states in question in some way expressing what we are to take to be an intuitive normative judgment about his well-being. Sumner stresses this alternative way of reading Griffin's case as an argument against simple hedonism: "On Bentham's version of the classical view, Freud seems plainly to have chosen the option which was worse for him. But that is a judgment few of us would join in making, and one which he himself would presumably have rejected."[12]

But remember the first characteristic danger of concrete cases which I noted earlier. The Freud case is complex, and its

very impurity makes it hard to isolate those of our reactions which reflect intuitive judgments about individual well-being. Freud was surely moved by other than merely his own well-being, given the grand nature of his intellectual ambition and self-conception and the therapeutic nature of his overall project. He may legitimately have preferred clear thinking, but on grounds other than a mere selfish concern with his own well-being. So this concrete case is inconclusive.

Considerations of the general sorts which are in play in this first case arise again in Griffin's more extended second objection, where Griffin and Sumner focus most of their discussion. So let us turn there. Here again is the entire relevant passage from Griffin:

> The trouble with [the hedonist] account is that we do seem to desire things other than states of mind, even independently of the states of mind they produce. This is the point that Robert Nozick has forcefully made with some science fiction. Imagine an experience machine programmed to give you any experience you want; it will stimulate your brain so that you think you are living the most ideal life, while all the while you float in a tank with electrodes in your brain. Would you plug in? "What else can matter to us", Nozick asks, "other than how our lives feel from the inside?" And he replies, surely rightly, that we also want to do certain things, to be certain things, and to be receptive to what there is in life beyond what humans make. The point does not need science fiction; there are plenty of examples in ordinary life. I certainly want control over my own fate. Even if you convince me that, as my personal despot, you would produce more desirable consciousness for me than I do for myself, I shall want to go on being my own master, at least as long as your record would not be much better than mine. I prefer, in important areas of my life, bitter truth to comfortable delusion. Even if I were surrounded by consummate actors able to give me sweet simulacra of love and affection, I should prefer the relatively bitter diet of their authentic reactions. And I should prefer it not because it would be morally better, or aesthetically better, or more noble, but because it would make for a better life for me to live. Perhaps some such preferences, looked at with a cold eye, will turn out to be of dubious rationality, but not all will. This fact presents a serious challenge to [the hedonist account.][13]

First point: Notice that for the most part this passage also simply observes that we desire, want, or prefer things other than our own hedonic tone. And still, by now on the next page, what we prefer is rejected by Griffin as a criteria for our well-being on "overwhelming" grounds.

There are, though, other elements in this passage. The third-last sentence claims that interfering factors – factors other than intuitive judgments of individual well-being – do not play a role in our reaction to the cases, or at least no ineliminable role, that the cases are suitably pure. The second-last sentence approaches the claim that some of the preferences in question are correct, though perhaps it does not quite get there, since the claim that something is not of dubious rationality may be merely the claim that it is not mad or incoherent and not quite the claim that it is correct in the relevant sense. And it is striking that though Griffin says that some of the preferences he's noted are not of dubious rationality, he does not tell us which ones. But my main point is that the claims in the second- and third-last sentences are controversial and partly normative claims about the concrete cases, and not parts of those cases.

There are many elements in these concrete cases which disturb our ability to factor out confounding elements. The experience machine is unfamiliar gadgetry which invokes our fear of the unfamiliar. It certainly is wildly unrealistic (Haslett, 1990). It involves a troubling irrevocability (Goldsworthy, 1992). And it seems to at least threaten risks of even hedonic harm that the corresponding actual life would not present.[14] Griffin deploys other concrete cases, from ''ordinary life'' and not from science fiction: the personal despot and the consummate actors. But of course these are not really all that realistic. Consummate actors would not realistically devote their lives to the project suggested. And even if they did, there would always be the realistic possibility that things might go radically amiss, in a way that would not threaten an ordinary life such as the actors are only simulating.

A second important cluster of confounding factors that influences our reaction to all these cases is that things other than our own well-being are in play in all of them, and matter to us. For instance, our lives have effects on other people which are quite significant, and which we care about, and which the judgments Griffin favors apparently ignore.[15] It is very hard to suspend consideration of those effects. It is hard to forget that one's loved ones would be hurt in Nozick's scenario, and that

one might blunder badly, through the ministrations of the consummate actors, in any attempt to help those one cares about.[16]

Mill's satisfied pig, who you will recall it is supposed to be worse to be than Socrates dissatisfied, is the ancestor of all these cases (Mill, 1979). But of course serious contemplation of becoming a pig invokes interfering factors of all the kinds we have traced, with a little species pride thrown in.

There is indeed yet another cluster of considerations which interfere with our ability to draw clean judgments from these cases, but it is perhaps best to consider these in light of the other general set of worries about the use of concrete cases which I noted earlier. As I said, it is important to consider a range of concrete cases all at once to form a properly balanced judgment about our complex normative intuitions.

Let me focus specifically on the intuitive significance of living in the truth. Griffin's cases do reveal a certain revulsion we feel not merely to being deceived, but also to false experience itself. The personal despot may more manipulate than deceive, and it is interesting that there Griffin retreats to a relatively weak claim about control over one's own fate, that he would prefer to be his own master as long as the benevolent personal despot did not have a much better record. Still, even just the truth of an experience involves more complexities than a single case or a few sets of cases can easily reveal.

We all grant that truth is the characteristic epistemic value of beliefs. So when other things are equal, of course true beliefs are better in some sense, though not necessarily a morally relevant sense partly constituting individual well-being. And of course true beliefs allow more certain action for one's goals, whatever they are. But the crucial question is whether living in the truth is of significance to well-being in itself and not for its various effects, for instance whether there are decisive commonsense intuitions which show that someone would be intuitively better-off if less happy but more in the truth. There is some suggestion of this in Griffin's claim about bitter truth. But our intuitions on these matters are complex and delicate, and not adequately revealed by our reaction to a single case.

Sumner's discussion uncovers some of the relevant complexities. To understand this discussion, it is useful to know that *his* view is that welfare consists of authentic happiness.[17] Happiness is supposed to be life-satisfaction, which has an affective component which involves experiencing the condition of your life as fulfilling or rewarding, and a cognitive component, which involves judging that your life is going well. Such happiness is supposed to be authentic when it is informed, and autonomous. But of course the crucial question is how to balance off authenticity and happiness. And Sumner has interesting things to say on this topic, our current concern.

Sumner suggests plausibly that Griffin's cases alone may suggest too rigid a view. Virtual reality machines might be something that we would enjoy in limited ways. Here it seems perhaps that individual preferences should guide us in judgments about well-being. Some people may like experience machines for some purposes and some may not. And there are even more difficult cases to consider. Sumner considers a mother asking a sergeant whether or not her soldier son died in agony, saying that she would rather be so unhappy that she would want to die than to delude herself. But then, after having been informed about the son's agonized death, she writes, "I can't write it. I wanted to know and got what I thought I wanted."[18]

In the end, in light of the various complexities he considers, Sumner proposes an account of the relevance of falsehood to well-being that rests on individual retrospective judgments. Consider

> the case of a woman who, for a while, lives in ignorant bliss with a faithless partner. Her endorsement of her life lacks information about his character and intentions. Is this information relevant? It is if her possessing it would undermine that endorsement. There are, therefore, two possibilities, which open up once she has been undeceived. One is that she re-evaluates how well her life was going … during the period of deception: "I thought everything was going well, but now I can see that it was all a farce." In that case, the discount rate she now imposes on her earlier assessment of her well-being determines how relevant the information was. The other possibility is that she does not care: "C'est la vie; at least he was charming and we had a lot of fun." Here the information turns out to have zero relevance, since that is the status she confers on it.[19]

One sort of objectivist account would hold that the truth of an experience matters to its contribution to well-being, period, whatever our individual belief and preference on the matter. That still leaves the problem of determining how much it counts, how much happiness is worth how much truth. But the main problem with such an account is that it is unintuitive to insist on the significance of truth despite individual preference to the contrary. Sumner's is a kind of desire-based account,[20] and appropriately and attractively rests the issue on individual preference. The relevant preferences cannot be revealed during the deception, but they can be afterwards, retrospectively. Sumner goes on to require that the revealed preferences be informed and authentic, and we might wonder about whether there is a determinate fact of the matter about what we would desire if fully-informed. But there are other worries also. We can so assess the state of those *terminally* duped only by reference to counterfactual enlightenment, and there are a range of different possible counterfactual enlightenments which might lead to different judgments.

But what's worse, and in the end decisive against Sumner's account, is that it is implausible to assume that actual (or hypothetical) retrospective judgments will be stable over time. On good days our now undeceived woman may be glad she was deceived and had a good time, and on bad days she may not be. And both attitudes may survive confrontation with full information. Even when such a judgment is stable, it may be an irrelevant accident that it is so. Sumner's retrospective account hence fails, at least in some cases.

Obvious objectivist accounts of well-being which invoke the significance of the truth of an experience fail intuitively, because they fail to take individual differences seriously. But obvious desire-based accounts which invoke truth also fail for the reasons just noted. Since there is no extant account which handles all extant cases relevant to assessing the role of truth to well-being in an obviously intuitive way, this class of cases hardly constitute decisive objections to hedonism which privilege competitor accounts over it. Since we do not know that all our intuitions about these cases can be made consistent and

coherent, we do not know from these cases that hedonism is false.

And notice that desire-based accounts such as Sumner's fail in this context in a revealing way, a way which reflects a general difficulty for some of the apparent concrete counterexamples to hedonism which we have been discussing. There is more than one perspective invoked by these examples, and it is not always obvious that one particular perspective is intuitively dominant, at least when a full range of cases is considered. The experience machine, the consummate actors, Socrates and the pig, and the deceived woman all involve a very unusual sort of dual or multiple perspective on things which makes them unrealistic or unusual, which hence perhaps directly undercuts the probity of the intuitions that they generate. But it also generates the important issue of whose perspective takes precedence, even whose corrected perspective. In the experience machine case, one makes a judgment regarding one's later state from an original perspective in which one feels discomfort about the possibility of deception, but such that later one will not feel that deception because in fact one will be deceived. It is not obvious that the first perspective has automatic normative authority over the first, and indeed it is not obvious that the view from the first perspective is not distorted by irrational discomforts. We can see this more clearly in Sumner's mother case. Or consider this case: You are amputating Ahab's leg without anesthetic, and he is screaming that he would rather die. Then he wakes up the next morning and thanks you. You feel better, of course. But why think that Ahab in the morning is the better authority on what the pain was worth? After all, he is not feeling it anymore, and it is safely in the past.[21]

We can evade differing temporal perspectives in these cases only by considering counterfactual judgments or actual judgments about hypothetical cases. But of course people's actual judgments about hypothetical cases may not reflect their judgments should things like that in fact become actual. When we ask real deceived people whether they would like to always know the truth, we can reasonably suspect their positive answer is a hope that they will find out that their illusory beliefs are

true. The relevance of that attitude is undercut if they turn out later to be miserable in light of the truth. And we noted earlier that there may be an indeterminacy regarding the judgment which a hypothetically fully-informed person would make. What's more, we just saw that even fully informed judgments may not be temporally stable.

We have seen that the standard concrete cases which apparently undercut hedonism are not sufficient, for a variety of reasons, to dismiss hedonism as uncompetitive. But, indeed, the issue of the relation of true beliefs to well-being does on balance suggest some intuitive grounds to prefer hedonism to its competitors. Let me turn now to that positive case.

Clearly, another important and relevant set of concrete cases probing the relationship of truth and well-being stems from the observation that, depending on how we factor out the effects on other people from the concrete cases we have been considering, we may be left with cases which are oddly reminiscent of some familiar metaphysical views which don't intuitively threaten well-being.[22] Presume that Berkeley is right about the world. It would seem intuitively that your well-being is as great as it would be if for instance Cartesian dualism were true, if all other relevant facts were unchanged. Presume that you feel good about your life because you have climbed a large mountain in Berkeley's world. Now ask what's the relevant difference between being in Berkeley's world and in Nozick's machine? Perhaps it is normatively relevant that other people experience things as you do in Berkeley's world but not in Nozick's. Still, to cleanly assure ourselves that we are not influenced by considerations of other people's well-being in making our judgments about such a case, we would have to put everybody affected relevantly by your life, say your whole family, into a similar machine. Now what's the difference? Perhaps you'd still worry that some people, or God, would see that the whole family was being manipulated, and indeed that some were engaged in that manipulation. But consider the coordinating role which God plays even on Berkeley's conception. Why is that any better? Or for that matter, consider the coordinating role of the galilean colorless world on some physicalist conceptions of

our color experience. These things seem analogous to the manipulation you would undergo in the experience machine, at least if enough people were placed with you in coordinated experience machines. And it should be that, if anything, actual manipulation and not merely people's perception of it is what matters, since those perceptions might be false, and we would not want to quickly grant that other people's false perceptions make us worse off in some intrinsic way. If people think you are deceived but in fact you are not, probably that shouldn't matter *directly* to your well-being, independent of its effects. Of course, philosophers often have cherished metaphysical views. Perhaps it is deeply significant to you whether idealism or physicalism is true, and if Berkeley's world were true you would judge your life much worse. But then that fact will distort your judgment about this case as a clue to the general significance of truth to individual well-being. We should consider rather the judgment of the non-philosopher who has no idiosyncratic commitments at risk in just this area.

The obvious parallels between these metaphysical cases and more ordinary deception cases suggests some support for the hedonist treatment of deception. And my positive case has another component. There is yet another set of related cases which we need to bear in mind to adequately judge the intuitive significance of the truth of experience for well-being. These, like the metaphysical cases, are related to cases we have reviewed which apparently tell against hedonism, yet they intuitively tell, on balance, against at least the dominant desire-based accounts and for something like hedonism. These cases indeed invoke a characteristic difficulty for desire-based accounts (Overvold, 1980, 1982). The true satisfaction of desires may make no difference in one's experience, and in a way which hence has no obvious intuitive relevance to one's well-being. One may have desires for something to happen to a stranger one meets on a train, and the fulfillment of those desires has nothing intuitive to do with one's well-being.[23] And many object to the thought that one's well-being can be affected by things that happen after one's death, even if some informed desire is fulfilled after death. The obvious way to modify desire-based accounts to evade

these difficulties is to build in an experience requirement, that one experience the satisfaction of the relevant desire. But of course that does not allow desire-based accounts to deploy some of their alleged intuitive advantage over hedonism in dealing with the truth of experience.[24]

Griffin, for reasons like this, doesn't build an experience requirement into his characteristic desire-based account, and grants that it is in some sense initially intuitive to insist that satisfaction of desire after death is irrelevant to well-being. But he argues that there is a slide between cases which are intuitively better for his view and these after-death cases, which should adjust our intuitions about the latter:

> Some of our aims are not fulfilled until we are dead; some, indeed, being desires for them, could not be. But is this so embarrassing after all? You might have a desire–it could be an informed one, I think–to have your achievements recognized and acknowledged. An enemy of yours might go around slandering you behind your back, successfully persuading everyone that you stole all your ideas, and they, to avoid unpleasantness, pretend in your presence to believe you. If that could make your life less good, then why could it not be made less good by his slandering you with the extra distance behind your back that death brings? You might well be willing to exert yourself, at risk of your life, to prevent these slanders being disseminated after your death. You might, with full eyes open, prefer that course to longer life with a ruined reputation after it. There seems nothing irrational in attaching this value to posthumous reputation. And the value being attached to it doesn't seem to be moral or aesthetic or any kind other than the value to be attached to the life as a life to be lived.[25]

There are various things to note about this paragraph which are related to problems with Griffin's cases against hedonism which we have already discussed. Again, his appeal is largely to our actual preferences, and once again there is a pretty weak claim to non-irrationality playing a crucial role. What is more, value attached to a life as a life to be lived is not obviously the same as well-being. It is also hard to control in our reactions to this case for the real risks for real unhappiness such slander during your life would imply. But my main point is that a slippery slope argument can be run in two directions. If we are confident that the satisfaction of desires after death does not matter to well-being, as many are, then Griffin's argument should con-

vince us that slandering us behind our backs does not in itself affect our well-being during life. What makes us first and misleadingly think there is a difference in the cases is simply their impurity, the fact that slander while we are alive creates all sorts of risks that we will suffer experienced effects even if in fact we would not ever suffer those effects.

This is not to say that there are no currents in our common sense which suggest the significance of what happens after death to well-being. But they are hardly decisive. The majority of philosophers who are not theologically inclined should be especially wary of distortions which have been introduced by the long historical focus on salvation. Aristotle perhaps introduced this topic into philosophy (Aristotle, 1985), and his own conclusion on the topic does not support a strong significance for what happens after death. It is probably also relevant that much of his discussion of the issue really concerns the difficulties which can overturn happiness before the very moment of death, and that some of his wording suggests literal causal effects on the dead which we do not take seriously.

Certainly, there is plenty in our tradition to suggest that what happens after death does not matter. If there is any text which can claim to be central to our tradition, Sophocles' *Oedipus the King* is such a text (Sophocles, 1984). It's final line is "[C]ount no man happy till he dies, free of pain at last."[26] It is also interesting that the final speech of the chorus, which concludes with that line, in fact suggests a link between the issues of effects on the dead and the general importance of deception and truth to well-being. The overall suggestion of the play, surely the most important and dramatic example in our literature of a life lived, for a while, in serious falsehood, is in accord with the suggestions of hedonism. Oedipus falls calamitously into a bad life, the play suggests, when he finds out the truth about himself. It contrasts his former greatness with the black sea of terror which has *now* overwhelmed him.[27] Immediately following the main revelation to Oedipus, the chorus says that the joy of his life is ground down to nothing.[28] The messenger who reports Jocasta's death and Oedipus' self-blinding in the next scene says of the pair:

> The joy they had so lately,
> the fortune of their old ancestral house
> was deep joy indeed. Now, in this one day,
> wailing, madness and doom, death, disgrace,
> all the griefs in the world that you can name,
> all are theirs forever.[29]

Don't we really agree that he was better off not knowing? Of course, this does not imply that Oedipus would not have been better off than he actually was even before finding out the horrible truth if his then false beliefs had been true. But the dominant suggestion of the play is certainly that living in falsehood, living in falsehood in the grandest and most terrible manner in our literature, is of quite limited significance to well-being. Otherwise there would be no calamitous fall for Oedipus upon his discovery of the truth.

At the very least, the relationship between truth and well-being is too intuitively complex for the concrete cases of deception we have surveyed to firmly establish that desire-based accounts or objectivist accounts are superior to hedonism. And indeed the great difficulties and complexities which they introduce into these accounts suggest some argumentative advantages for hedonism, which at least gives a simple and definite answer that is not obviously false for all the cases, and is supported by some intuitions. Still, this is not to deny that the concrete cases which are customarily deployed against hedonism reflect deep and general intuitive currents which are worthy of respect and further consideration, and that a deeper consideration of trade-offs between happiness and truth is required. We will return to these issues in a more general and abstract way in Section 4. While the concrete cases which are customarily deployed against hedonism are not alone adequate to privilege extant competitors over it, we will try to probe the deeper and more important general currents beneath those cases later.

3

While our discussion has already involved some concrete cases which concern objectivist theorists, our primary focus so far

has been largely on the dominant desire-based accounts. Contemporary arguments for objectivist accounts of well-being over hedonism, at least to the degree that they are distinct from the alleged counterexamples which we have already discussed, seem to rest less on particular concrete cases, and perhaps more on the general intuitive considerations which we will discuss in the next section. Still, let me say a little about some relevant and relatively concrete cases which may seem to specially favor objectivism.

Objectivist accounts have a variety of forms, but all specify individual well-being or good without regard to individual psychological attitudes. Perhaps there is a set of basic and objective human needs, say paradigmatically food or health, which we all have independent of any subjective desires or pleasures which we feel. These needs may define our well-being. Or perhaps there is an objective form of human flourishing of a more traditional aristotelian and perfectionist sort. Or perhaps, as Rawls suggests in a *Theory of Justice*, we should compare the well-being of individual lives by comparing amounts of "primary goods" possessed by the individuals in question, where primary goods are, more or less, universal instrumental goods, goods which one wants as a means of pursuing one's projects whatever those subjective projects turn out to be (Rawls, 1971).

Consider the first and second forms of objectivism, needs-based and perfectionist accounts. There are such accounts which make room for the significance of truth. The most obvious forms have intuitive difficulties we have already noted, in failing to allow for the proper treatment of cases in which individuals fail to care about the truth of their experience. But the health of those who don't care about it and receive no pain from ill-health is another obvious reasonably concrete case which may intuitively support objectivism of these sorts against hedonism or desire-based accounts.

But such a case is double-edged, since there are also related intuitive difficulties for objectivist accounts of well-being which, unlike hedonism or desire-based accounts, are not rooted in individual psychological states. While it is clear that health has intuitive relevance to well-being, that is most obvious when we

care about it, and hence when its value is recognized by desire-based accounts or other mental state accounts like hedonism. Objectivist accounts, to the degree that they do invoke features of people's lives which they do not care about, are not particularly intuitive. Unlike both hedonist and desire-satisfaction accounts of well-being, they can be criticized as intuitively insufficiently subjectivist, as paying too little attention to relevant features of individual psychologies.[30]

Another relevant point about health is this: Many things have intuitively healthy states, for instance rhododendrons and corporations, and even cancers and viruses. But it is probably the dominant intuition that it is in particular things with psychologies which are of genuine normative significance. If life matters for its own sake, it is hard to see why all life does not matter for its own sake.[31] And it is certainly hard to balance the health of different organisms.

But the main point is that it is not obvious to common sense that health is important to an individual's well-being when at least their rational individual preference suggests otherwise. Other things may matter more even selfishly. Imagine spending some health on salvation or artistic, athletic, or intellectual achievement.

Similar intuitive difficulties trouble the third-type of objectivist account. Universal means, of the type of Rawls' primary goods, involve elements which because of individual peculiarities of projects or motivation may bear insufficient relation to the things someone cares about, and hence are not intuitively relevant to their well-being. What is more, some individuals, for instance those with certain sorts of handicaps, are not able to transform such objective "goods" into intuitive well-being as efficiently as others. They hence seem subject to discrimination by such accounts, and indeed in a way which suggests that there must be some deeper phenomena against which we intuitively assess individual well-being and hence their relative discrimination. This is not to deny that institutions of certain particular kinds should distribute some generally helpful good which it is their special concern to distribute, for instance money or food or medical treatment, without regard to individual peculiarities.

But, as Griffin has argued, there is plausibly a rationale explaining why these generally helpful goods are good which lies in their aiding some human well-being of a deeper sort.[32] It is by reference to that deeper well-being that we can see that wealth is a good. But it is also by reference to that deeper sort of well-being that we can argue, as for instance Plato argues in the *Laws*, that too much wealth is destructive of well-being, or that certain individuals for special reasons should receive less or more of some particular generally helpful good, or instead some alternative good which is specially helpful to them. Levels of generally helpful goods are not intuitively sensitive enough to determine the well-being of those who are unusual in various ways.

One complication: Hedonism itself is in a recognizable sense an objectivist view. But still, our pain which we do not care much about in the specific sense involving desire is yet at least recognizably relevant to our own selfish good.

We have so far identified no clear intuitive support for objectivism over hedonism. Indeed, the reverse may be true. There are two final classes of relatively concrete intuitive objections to hedonism we have yet to consider, which are not specifically identified with either objectivist nor desire-based accounts. Perhaps they simply suggest a hedonism more sophisticated and qualified than my own, but in any case we should consider them.

First, it has long been held that certain pleasures or desires should not count normatively because of their sources. Foolish happiness, which we have already discussed, may be an instance of this, but there are others which are more characteristic of this set of objections. It may seem intuitive to claim that the pleasures of sadism or successful murder have no positive normative significance at all. And if the oppressed feel no pain about something, that may be just a sign of their deeper and more insidious psychological oppression (Sen, 1987). Second, it may seem that the pain of the evil is a good thing, and the happiness of the evil is bad thing, even independent of its particular source for them. Kant famously says at the beginning of the first section of the *Groundwork*: "[A]n impartial rational spectator

can take no delight in seeing the uninterrupted prosperity of a being graced with no feature of a pure and good will, so that a good will seems to constitute the indispensable condition even of worthiness to be happy (Kant, 1996a).''

My first point in response is that such cases are not all on balance intuitively troubling against simple hedonist conceptions, or even against contrary simple desire-based or objectivist conceptions, of *individual well-being*. Even intuitive retributionist conceptions of punishment must suppose that bad things happen to the evil-doer who is punished, that the punishment is bad for them. Otherwise it would not be a punishment. And the sadist is after all intuitively one who finds their own good in the suffering of another. That is part of what makes sadism so troubling and abhorrent. Note also, in the quote from Kant, that his point is that prosperity and happiness themselves are what those without pure will do not deserve. Pain is bad for the evil as for the good, and pleasure is good for them equally. If not, these things would not intuitively constitute punishments and rewards which might be justly or unjustly distributed.

Still, it might be that non-autonomous pleasures, in the sense we have already noted in our discussion of Sumner, provide decent concrete cases which tell intuitively against hedonic conceptions of well-being. But note also that these cases provide intuitive difficulties for any familiar account. Desire-based accounts obviously suffer analogous difficulties, and no intuitively plausible objectivist account can simply ignore the intuitive relevance of happiness to the well-being even of the deeply psychologically oppressed. Oppressed as they are, would they be better off being miserable also?

And yet, there is a sense in which the oppression of someone may play an intuitively negative role in their well-being which is not easily captured by hedonist accounts, analogous to the intuitive negative role of false experience which Sumner and Griffin do succeed in invoking. And there is in any case a recognizable intuitive pull to the claim that the pleasure of the evil should not count normatively as good even if it does count intuitively towards their abhorrent well-being. There is, we

should admit, some contrast between one's well-being and goodness *tout court*. Perhaps the pleasures of the guilty or sadistic ought to count against the overall value of a state of affairs, and not for it at all. The pains of the guilty may not be bad from the point of view of the universe even if they are bad for the evil-doer who suffers.

To address these matters properly, we will need to turn now from a piecemeal consideration of concrete cases relevant to individual well-being to a more general consideration of intuitive features of competing conceptions of the good. The hedonic tradition is one version of a plausible insistence that overall good must be ultimately rooted in the well-being and good of the sentient. But issues about punishment and reward remind us of the possible difference between conceptions of individual well-being which have been our main concern in this section, and conceptions of the good. The latter will be our primary focus in the next section. And of course we have also seen throughout this and the preceding section that single concrete cases are indecisive even for issues about intuitive individual well-being. We must, it seems, consider things more generally and abstractly if we are to make adequate progress in developing our understanding of commonsense commitments about well-being or the good.

4

The methodological difficulties presented by reliance on concrete intuitive cases suggest that it would be better to deploy abstract cases. This can at once assure that impurities are minimized, and that a relatively wide range of particular concrete instances are considered simultaneously at least by implication. The best discussion of well-being or the good with this form that I know, which in fact incorporates most of the various important general factors suggested by the concrete cases we surveyed in the last two sections but in a purer and more systematic way, is 8 pages of W.D. Ross' *The right and the good*.[33] This discussion is reminiscent of and probably modeled

on Plato's excellent discussion in the *Philebus*, but applied to a wider range of goods.

Ross argues that four things are intrinsically good: pleasure, virtuous disposition and action, knowledge, and the proper apportion of pleasure and pain to the virtuous and vicious.[34] He does so by deploying a characteristic pattern of argument to each of the relevant cases. Here is the case against mere hedonism and for the intrinsic good of virtuous disposition and action:

> And if any one is inclined … to think that, say, pleasure alone is intrinsically good, it seems to me enough to ask the question whether, of two states of the universe holding equal amounts of pleasure, we should really think no better of one in which the actions and dispositions of all the persons in it were thoroughly virtuous than of one in which they were highly vicious. Most hedonists would shrink from giving the plainly false answer which their theory clearly requires.[35]

As honest hedonists we should admit that, if the profile of hedonic tone present in two situations is identical, and if the probabilities of future profiles of hedonic tone associated with those situations is likewise identical, and if in one situation all actions and dispositions are vicious and in the other virtuous, then we are committed to the answer Ross thinks plainly false. Nevertheless, we have a reasonable response to his objection. Arguments with this general pattern, which may seem to support against mere hedonism the basic normative significance of virtue or knowledge or the retributory distribution of pleasure and pain, are misleading. The paired cases under consideration involve equalized basic value in the sense introduced in section 1, I claim. Therefore they magnify the intuitive significance of any other sorts of value which they contain. But of course this need not be basic value.

Indeed, Ross misses some kinds of value which are delivered by his test and which are in no intuitive sense directly morally relevant. There are intuitive characteristic excellences with a certain rough degree of specificity for all sorts of human projects, activities, artifacts, and institutions, and also for entities and organs created by mechanisms of natural selection or which mimic natural selection. Equalized cases such as Ross

deploys can magnify the intuitive significance of such intuitive excellences and goods until they seem to be of basic normative significance. But they are not. Or at the very least, not all of them are.

Some instances of this phenomenon are the goods beyond pleasure which Ross suggests are also intrinsic: (i) As I have already indicated, consequentialism can deploy basic value to deliver the derived normative significance of virtues and the right. And considerations of equitable distribution, and not merely the amount of basic value they contain, may plausibly constrain our sense of the relative value of overall states of affairs. (ii) The truth of beliefs is an obvious epistemic value, and there is perhaps a further epistemic value of knowledge beyond mere true belief.

And there are other cases which Ross' method shows to be significant in some sense, which he ignores, and which cannot plausibly be incorporated into any overall account of basic normative value in all of their instances. They include (i) the health of all organisms and even of organs and cells,[36] (ii) the satisfaction of at least most desires or whims, (iii) the excellence of forms of practice of almost any sort, (iv) the significance of works of art according to any set of coherent standards. All these things, all members of each class, can be made to seem significant by Ross' method. If other sorts of value are otherwise equal, why wouldn't it be better for there to be a healthy rhododendron, a satisfied whim, a flourishing corporation, or better intonation in the woodwinds? Of course, these new sorts of value will conflict in real cases, which set the health of one organ or success of one institution against another. But equalized cases of the sort which Ross deploys suggest that they have some sort of value nonetheless. But it need not be basic normative value, nor even value of any ethical or political significance.

We cannot tell by the method of equalized cases the nature of the value in question, whether for instance the value has basic significance for ethics, or whether indeed it reflects forms of assessment which can be justified at all. And yet we saw that concrete cases alone are deceptive. So what are we to do? We

might retreat to the claim that commonsense intuition can tell us nothing of significance. But I believe that we can do somewhat better. We did somewhat better even in the last two sections, for instance in our treatment of truth and unrecognized satisfactions, where we saw enough pattern in our intuitions about relatively concrete cases to suggest some intuitive advantage for hedonism. But clearly some further development of the method of abstract cases is desirable.

As a first step, we should develop Ross' abstract intuition tests so that they treat a wider range of cases than simple equalized pairs. If we consider abstract trade-off cases, in which we force our intuition to balance the significance of more of one good and less of another, we may still be unable to fully assess the intuitive nature of the values involved, but we can get a better sense of their intuitive significance and depth, of their overall importance when the chips are down. Important and dominant good at least points in the direction of the sort of basic value which we seek.

Ross himself makes some steps in this direction.[37] But, since his treatment is the best intuitionist treatment of these issues which we possess, it is unfortunate that his development of the topic is incomplete and in some particulars not very convincing. It would indeed be a very large task to treat abstract trade-off cases properly. Still, we can gain some insight from Ross' attempt.

What I am calling trade-off cases are pairs of situations in which it is specified that, holding all other sorts of value constant (and all attendant counterfactuals and probabilities relevantly identical and hence their contribution to intuitive value identical), one situation is clearly worse in respect of one sort of value but better in respect of a second, while the second situation is clearly better in respect of the first sort of value and worse in respect of the second. To prevent various concrete irrelevancies from influencing our intuition, and so that these pairs capture at once a range of instances, we should at least as a first stab articulate the pairs in something like the pure and abstract manner which Ross deployed. We should develop abstract trade-off cases.

While it is important that the cases we consider be abstract and pure, it may be that there are some additional concrete conditions which are useful in avoiding obviously misleading currents in our intuition when dealing with the significance of certain goods like pleasure. For instance, perhaps it is important that the pleasure in question not be our own, because our normative intuitions are not cleanly engaged in cases where our own pleasure is at issue, partly because of natural selfishness and partly because of the desire to avoid obvious selfishness. But such concrete additions may be themselves dangerous and misleading, and perhaps sufficient abstraction can also guard against most misleading elements in our responses. So let us attempt this method.

To narrow our treatment to the cases which seem most significant for our purposes, I will consider trade-off cases which range across the goods deployed in the alleged counterexamples to hedonism discussed in the last two sections and also the goods which Ross' own discussion suggests are intrinsic goods. Ross fails to isolate the intuitive intrinsic value of health which some objectivists may favor, and also that of unrecognized desire-satisfaction. But as I said he does isolate virtuous action, knowledge (implying the truth of belief), and the concordance of pleasure and virtue. He also suggests, in a manner apparently supported by his test, that pleasure is an intrinsic good somewhat unlike the others. It is an intrinsic good only when it has appropriate causes, when it is not vicious pleasure. We should consider this qualification on hedonism as well.

There are two relevant kinds of trade-off cases, those which compare goods within the abstract categories just noted and those which make cross-category comparisons. Consideration of the first sort of pairs may sometimes help put the second sort in the proper light. Consider health. Any genuine foundational role for health requires some account of how the health of one cell or organism or organ or eco-system trades against another. Nevertheless, let me focus on the crucial cross-category comparisons, in the first instance the trading of pleasure for health.

Here we can be guided by our discussion in the last section, which was at least implicitly abstract. Imagine the relevantly

abstract pair of situations. The kind of health which is not eventually positively reflected in the hedonic tone of the creatures involved, the kind our pair introduces, seems for that very reason of lesser significance. Intuitively, the health of the sentient is the health that matters, and its worth is at least reflected in its effects on that sentience. Intuition cedes, it seems, the dominance of pleasure over health in this abstract trade-off case. So Ross was right to ignore these trade-offs. The method we are attempting works for these cases, but doesn't reveal an interesting new basic good.

What of unrecognized desire-satisfaction? Here, as we saw at least implicitly in Section 2, there are no firm commonsense intuitions which resolve these issues abstractly and at once in favor of competitors to hedonism. Rather, there are conflicting intuitions, which for instance can slide up or down the continuum between unrecognized satisfactions received after death and those received during life. There is not a uniform intuition abstractly rooted. Concrete details matter in conflicting ways. I believe that we also saw on balance in section 2 that cases of this general sort, indeed of a more general sort which includes the significance of true belief to well-being, favor hedonism. But the methodology of abstract cases itself apparently fails to achieve adequate resolution in this application.

So far abstract trade-off cases have yielded little. But there is a still more refined version of the method we might attempt. It might be that relatively general factors found in the concrete cases which favor hedonism regarding the significance of true belief and unrecognized satisfaction in the last two sections can somehow be parlayed into specially relevant general conditions, which allow us to specify extra but still general conditions on what we might call "relatively abstract" trade-off cases. The conditions might allow our intuition some determinate purchase, yet be conditions which we can see to be appropriately revealing and not inappropriately misleading.

Still, we should not be confident of that possibility for trade-offs involving unrecognized satisfaction, since the factors to which I appealed in section 2 do not seem to be of the proper form to allow it. We rather relied, for instance, on the firmness

of intuitions about well-being after death and abstract similarities with other cases where intuitions were less firm. So again, the fact that Ross ignored these cases may seem hence reasonable. They do not seem good instances for his method or anything in its rough vicinity.

Still, the issue of unrecognized satisfaction is closely related to that of true belief, and this we can treat by reference to Ross' explicit discussion of that issue. So perhaps this provides an opening for something like Ross' method. Consider knowledge.

Here the first sort of abstract trade-off tests, which probe trade-offs within a single category of good, are quite interesting. Ross attempts to develop an account of how various forms of knowledge and true belief are properly to be traded against other forms, and while it is perhaps not convincing, we should be sensitive to the obvious difficulties which he faced. To work things out properly we would need for instance to consider the relative weight of truth as opposed to the third condition required for knowledge. I find it hard to believe that the presence or absence of false barns in one's vicinity, which consensus now holds relevant to the third condition, could plausibly be held relevant to one's well-being, but still perhaps something beyond the truth of one's belief might be intuitively so relevant. Ross considers such factors as the degree of match of conviction and evidence and also the generality of the content of the beliefs in question. There are many difficulties, and I commend his discussion to your attention, though it clearly requires further development.

But, however that development goes, the crucial cases for our purposes are relatively abstract trade-offs across categories of goods, of pleasure on the one hand and life in the truth on the other. How much suffering is more true belief worth? My own even fully abstract intuition is that true belief bought for suffering is of no deep and intrinsic worth to well-being, and is only good in an obvious epistemic sense of no intrinsic moral significance. I have a firm intuition reaction to even a fully abstract trade-off case of this sort. But perhaps that is idiosyncratic. Still, clearly the fully abstract trade-off does not count against hedonism in any definite way. We saw at least

that in section 2. Ross himself finds this general issue too difficult, and so retreats to the claim that the desire for knowledge is a virtue and that is hence why knowledge is of real but hence somewhat indirect significance.[38]

But our immediate concern is whether *relatively* abstract trade-off cases can be developed in a manner which provides intuitive resolution of this trade-off across categories. Still, the situation with true belief seems no better than that with unrecognized satisfaction. It is not obvious that the concrete cases which on balance support hedonism over competitors in this trade-off, as we saw section 2, have general features which can be appropriately imported into relatively abstract trade-off cases. We seem left in this case with the piecemeal method of that section.

Perhaps the crucial and most troubling cases for hedonism involves the foundational significance of virtuous action, the just commensuration of pleasure to virtue, and the related concern that vicious pleasures should not count normatively. Some concerns about a foundational significance for these things are in effect objections to consequentialism. Some are concerns about the value of overall states of affairs which include pleasures and pains as parts. They hence seem not to be concerns about what I have called the basic normative value of pleasure per se. Nevertheless, we can get some grip on the intuitive relevance of these things to basic value by a consideration of abstract trade-off cases, or at least by a consideration of relatively abstract trade-off cases. It is less than first may meet the eye.

If we consider a pair of situations, with the attendant counterfactuals and probabilities specified as relevantly identical, such that the first clearly involves more suffering and less pleasure but also more virtuous action (and perhaps knowledge) and also more intuitively commensurate suffering for the intuitively guilty, my own honest but not firm intuition is that the first is worse overall than the second. Still, it does seem plausible to insist that some of the relatively concrete details will matter if we are to establish any firm and clear intuitive reaction. If we specify that it is the vicious only who suffer

more, or that the suffering replaces a vicious pleasure, then my hedonist reaction is undercut. But if on the other hand we specify that it is the innocent who suffer more, my intuition is stronger still. The intuition in question is even stronger if the sufferer is one of the worst-off. And indeed if the "guilty" or "vicious" person is the worst off, this undercuts even my weak intuition that perhaps the guilty or vicious should suffer more, since it seems in the scenario in question that they have already suffered enough to count as appropriately innocent and should properly grab for what they can get.

Without some relatively concrete specificity, fully abstract trade-off cases are apparently indecisive in this instance. But in this situation I claim that at least a relatively abstract trade-off case is available and effective. There are relatively abstract features which seem relevant and can resolve our intuition. What is relevant is whether the extra suffering is for the relatively worse off.

But not everyone will agree. The issue really becomes whether this is the proper relatively concrete specificity to introduce into these trade-offs. This focus on the suffering of the worst off may seem to some an arbitrary and misleading concretizing of the pairs, because if we focus on extra suffering for the well-off and especially the well-off guilty or vicious, then we would have other intuitions.

I believe that it is possible to argue that a focus on the suffering of the worst off is not arbitrary here. I have argued elsewhere, following Rawls and Scanlon, that it is the appropriate way to consider any complex tradeoff of well-being and reflects the proper means of summing value into the overall value of states of affairs.[39] When we cannot form an overall intuitive judgment about such a trade-off pair described with full abstraction, then consideration of the worst off is the appropriate relative concretization to make. In this case, though not in the case of life in the truth, an effective and appropriate relatively abstract trade-off case is available, and it supports hedonism. That is because the vicious and the worst off may not be the same individual, so methods of balancing the interests of the worst off versus other moral interests become

most clearly relevant. A relative concretization of this sort may also allow us some purchase even on the relative general importance of truth and suffering, but not perhaps in the most characteristic and telling location to resolve that particular issue, where the suffering and the deceived are the same person. So in that case we are forced back to the intuitional method of the previous sections. Still, of course I cannot properly insist here on the probity of my other arguments in other contexts.

What is relevant here is a general and methodological point: The proper development of the abstract method of this section requires an understanding of the general and intuitive theoretical rationales for intuitive responses, and not just those particular intuitions themselves.

In the case at hand, as I have said, I believe such a rationale is available to us. Kant says that ''one has never heard of anyone who was sentenced to death for murder complaining that he was dealt with too severely and therefore wronged; everyone would laugh in his face if he said this (Kant, 1996b).'' But let me vary the figure. We might address a relatively badly-off suffering person and suggest that they might have been better off, but then the world would have needed to include more lying or ignorance or a more striking mismatch between virtue and happiness, even though for some reason all that would not in fact have made anyone else suffer more. If we concluded ''Things really are better this way, though you must regrettably suffer more'', then I think we should reasonably and intuitively feel very uncomfortable about that. Frankly, it would seem appropriate if they did worse than chuckle.

But a properly full treatment of this case would need to explain why this isn't merely rhetoric, why it is appropriate to make the relatively abstract trade-off cases assume this particular form. While I believe that one is available, that is of course a different story.

Absent a consensus on such a rationale, intuitions will differ and cannot resolve the issue. That is the crucial point of this section. But I also think a comparison with what Ross says puts my own intuitions, supported by a rationale or not, in at least a relatively good light:

I think … that pleasure is definitely inferior in value to virtue and knowledge… Most people are convinced that human life is in itself something more valuable than animal life, though it seems highly probable that the lives of many animals contain a greater balance of pleasure over pain than the lives of many human beings. Most people would accept Mill's dictum that 'it is better to be a human being dissatisfied than a pig satisfied'… Many people whose opinion deserves the greatest respect have undoubtedly thought that the promotion of the general happiness was the highest possible ideal. But the happy state of the human race which they aimed at producing was such a state as the progress of civilization naturally leads us to look forward to, a state much of whose pleasantness would spring from such things as the practice of virtue … [I]f they thought the state of maximum happiness would be one whose happiness sprang from such things as the indulgence of cruelty, the light-hearted adoption of ill-grounded opinions, and enjoyment of the ugly, they would immediately reject such an ideal… [What] amount of pleasure is precisely equal in value to a given amount of virtue? … [I]t seems to me much more likely that *no* amount of pleasure is equal to any amount of virtue.[40]

Some analytical remarks about this paragraph. From Ross' sensitive general method, we first get a retreat to the single and misleading concrete case of the contented pig, which is not supported (like relative concretization involving the worst off) by a suitable rationale. It *is* so supported in Aristotle, but Ross lacks his account of a species-specific human *telos*. Then we get an implausible prediction about the opinions of others, citing grounds for their rejection of pleasures which are to the contrary embraced by many post-modern fans of *juissance*. And we end worst of all, with a wildly counterintuitive and even offensive general claim. How *dare* we say to the suffering that no amount of their suffering would make the world worse when traded for a little more virtue somewhere, or a bit more general knowledge, or a bit more suffering for the intuitively guilty?

But, as I have said, Ross does provide the best and most thoughtful anti-hedonist discussion of the good rooted in intuition and couched at the proper level of generality and purity.[41] It just is not good enough.

We can properly conclude three things, in decreasing order of certainty. First, there are no extant competitor conceptions of the good or well-being which possess a definite overall

intuitive advantage over hedonism. This is not grounds, of course, to stop trying for something better than all current competitors. But the competitors span the three types in a natural characterization of possible conceptions of the good or well-being. And until we have such a coherent account of all our intuitions, we have no reason to believe it to be possible. Hence we come to my second conclusion. Many ethicists believe that they *know*, on grounds of intuition, that hedonism is false. But they know no such thing. Third, at least if the rationale for a specific appropriate concretization of relatively abstract trade-off cases which I have promised in this section is available, but to a large degree even if it is not, hedonism is better supported by our overall intuition than its competitors.[42]

NOTES

[1] Mendola, 1990a, argues that hedonism is true independent of intuition.
[2] Mendola, 1990a.
[3] It is a stretch to call Sumner a hedonist. We will return to his view.
[4] Edwards. Sidgwick.
[5] Sumner, 102–103.
[6] Mendola, 1990b, forthcoming.
[7] He distinguishes between objective accounts and perfectionist accounts, which I will treat together.
[8] Sumner, 92.
[9] Griffin.
[10] Griffin, 8.
[11] Griffin, 10.
[12] Sumner, 92.
[13] Griffin, 9.
[14] Sumner, 94–95.
[15] Haslett.
[16] I presume that some of these might not be among the consummate actors, or that one might even care about some of the consummate actors regardless of their lies.
[17] Sumner, 138–183.
[18] Sumner, 97.
[19] Sumner, 160–161.
[20] Though closer to classical hedonism than most.
[21] I owe this case to a public lecture by Thomas Schelling.
[22] Carson, 51–53.

[23] This case is attributed by Sumner to Parfit.
[24] It also requires that we develop some way of balancing experienced satisfaction of a desire accompanied by unknown frustration which, if it is not to collapse into hedonism, requires adjudication of some of the balancing difficulties that Sumner faced.
[25] Griffin, 23.
[26] Sophocles, 251.
[27] Sophocles, 251.
[28] Sophocles, 234.
[29] Sophocles, 237.
[30] Sumner, 45–80.
[31] Sumner makes similar points.
[32] Griffin, 40–55.
[33] Ross, 1930, 134–141.
[34] There are certain qualifications of his claims about pleasure and pain, to which we will return.
[35] Ross, 1930, 134.
[36] Indeed he seems to explicitly discount the significance of disease on 134.
[37] Ross, 1930, 142–154.
[38] Ross, 1930, 151–152.
[39] Mendola, 1990b.
[40] Ross, 1930, 149–150.
[41] This is the guts of Ross' discussion of these trade-offs. But he makes some other arguments, and I commend his overall discussion to your attention.
[42] Thanks for comments from Tom Carson, Mark van Roojen, and a referee for this journal.

REFERENCES

Aristotle (1985): *Nicomachean Ethics*, Indianapolis: Hackett, Book I, Chapters 9–10.
Casell, E. (1982): 'The Nature of Suffering and the Goals of Medicine', *New England Journal of Medicine* 306, 11.
Carson, T.L. (2000): *Value and the Good Life*, South Bend: University of Notre Dame Press.
Edwards, R.B. (1979): *Pleasures and Pains*, Ithaca: Cornell University Press.
Griffin, J. (1986): *Well-Being*, Oxford: Clarendon Press.
Goldsworthy, J. (1992): 'Well-Being and Value', *Utilitas* 4, 1–26, 18–20.
Haslett, D.W. (1990): What is Utility?, *Economics and Philosophy* 6, 65–94.
Kant, I. (1996a): '*Groundwork of the Metaphysics of Morals*', in M.J. Gregor (trans. and ed.), *Practical Philosophy*, Cambridge: Cambridge University Press, 37–108, p. 49.
Kant, I. (1996b): '*The Metaphysics of Morals*', in Gregor, 353–603, p. 475.

Mendola, J. (1990a): 'Objective Value and Subjective States', *Philosophy and Phenomenological Research* 50, 695–713.
Mendola, J. (1990b): 'An Ordinal Modification of Classical Utilitarianism', *Erkenntnis* 33, 73–88.
Mendola, J. (forthcoming): 'Consequences, Group Acts, and Trolleys', *Pacific Philosophical Quarterly*.
Mill, J.S. (1979): *Utilitarianism*, Indianapolis: Hackett, pp. 7–9.
Overvold, M. (1980): 'Self-Interest and the Concept of Self-Sacrifice', *Canadian Journal of Philosophy* 10, 105–118.
Overvold, M. (1982): 'Self-Interest and Getting What You Want', in H. Miller and W. Williams (eds.), *The Limits of Utilitarianism*, Minneapolis: University of Minnesota Press, pp. 186–194.
Parfit, D. (1984): *Reasons and Persons*, Oxford: Oxford University Press, p. 4.
Rawls, J. (1971): *A Theory of Justice*, Cambridge: Harvard University Press.
Ross, W.D. (1930): *The Right and the Good*, Oxford: Oxford University Press.
Ross, W.D. (1939): *Foundations of Ethics*, Oxford: Oxford University Press, pp. 252–310.
Sen, A. (1987): *On Ethics and Economics*, Oxford: Basil Blackwell, pp. 45–46.
Scanlon, T.M. (1998): *What We Owe to Each Other*, Cambridge: Harvard University Press (Chapter 3).
Sidgwick, H. (1907): *The Methods of Ethics*, 7th edn., London: Macmillan.
Sophocles (1984): *Oedipus the King* (translated by R. Fagles), in *The Three Theban Plays*, Harmondsworth: Penguin Books.
Sumner, L.W. (1996): *Welfare, Happiness and Ethics*, Oxford: Clarendon Press, p. 83.

Department of Philosophy
University of Nebraska – Lincoln
1010 Oldfather Hall
Lincoln, NE 68588-0321
USA
E-mail:jmendola@unlserve.unl.edu