ALASTAIR NORCROSS

Comparing Harms: Headaches and Human Lives

Consequentialists are sometimes unsettled by the following kind of example: a vast number of people are experiencing fairly minor headaches, which will continue unabated for another hour, unless an innocent person is killed, in which case they will cease immediately. There is no other way to avoid the headaches. Can we permissibly kill that innocent person in order to avoid the vast number of headaches? For a consequentialist, the answer to that question depends on the relative values of the world with the headaches but without the premature death, and the world without the headaches but with the premature death. If the latter world is at least as good as the former, it is permissible to kill the innocent. Furthermore, if the all-things-considered values of the worlds are comparable, and if a world with more headaches is, *ceteris paribus*, worse than a world with fewer, it is reasonable to suppose that a world with a vast (but finite) number of headaches could be worse than a world that differs from it only in lacking those headaches and containing one more premature death. In short, there is some finite number of headaches, such that it is permissible to kill an innocent person to avoid them. Call this claim *life for headaches*. Many people balk at *life for headaches*. In fact, many people think that there is no number of people such that it is permissible to kill one person to save that number the pain of a fairly minor headache. Deontologists might think this, because they endorse what Scheffler calls "agent-centered restrictions." Such restrictions forbid certain kinds of action, even when their results are at least as good as all alternatives. Thus, a deontologist might agree that the world without the headaches but with the premature death is

Earlier versions of parts of this article were presented at Texas A & M University, the University of Mississippi, the American Philosophical Association Pacific Division meeting, and the North Texas Philosophical Association. I thank audiences at all these events for helpful comments. The article has also benefited from the comments of Greg Bayer, Doug Ehring, David Hausman, Mark Heller, Derek Parfit, and Steve Sverdlik.

better all things considered than the world with the headaches but without the premature death, and yet maintain that it is impermissible to kill the person in order to avoid the headaches.[1] A consequentialist, however, who agrees with this ranking of the two worlds must also claim that it is permissible to kill the innocent person.[2]

I

In order to avoid *life for headaches*, therefore, a consequentialist must deny that a world with a vast (but finite) number of headaches could be worse than a world that differs from it only in lacking those headaches and containing one more premature death.[3] One reason for such a denial can be stated as follows:

> *Incomparable:* The state of affairs constituted by *any* number of fairly minor headaches is incomparable with the premature death of an innocent person.

To say that a state of affairs (or option) A is incomparable with a state of affairs (or option) B is to say that (i) it is false that A is better than B, and (ii) it is false that B is better than A, and (iii) it is false that A and B are of equal (or roughly equal) value.[4] There are at least two reasons why a consequentialist who wishes to avoid *life for headaches* should not be satisfied with *Incomparable*. First, a consequentialist who denies that it is permissible to kill an innocent person in order to avoid a vast number

1. Of course, it is one thing to *claim* that such a deontological restriction exists, it is quite another to *argue* for its existence. For an attempt to do the latter, see Judith Thomson, *The Realm of Rights* (Cambridge, Mass.: Harvard University Press, 1990), Chap. 6. For an argument that Thomson's approach does not succeed, see Alastair Norcross, "Rights Violations and Distributive Constraints: Three Scenarios," *Pacific Philosophical Quarterly* (June 1995): 159–67.

2. This assumes that agent-neutrality is an essential feature of consequentialism. This is widely accepted. However, I do not wish to go into the question of whether there could be a theory that endorsed agent-relativity and is still distinctively consequentialist in nature. This article is addressed to all those consequentialists, which may be all the consequentialists there are, who endorse agent-neutrality.

3. It is not clear what kind of possibility is required to satisfy those who find *life for headaches* so counterintuitive. Some may be content with nomological possibility, others may demand metaphysical possibility. For the purposes of this article, it does not matter.

4. Some philosophers equate incomparability with rough equality. This may involve a conflation of the notions of incomparability and of vagueness. I believe it avoids confusion to use a notion of incomparability that excludes rough equality as well as equality.

of headaches is most likely to do so because she thinks it *worse* to kill than to allow the headaches, not because she thinks the two options are incomparable. Thus, she will most likely also claim that it is permissible, in fact obligatory, to allow the headaches to continue, if the only alternative involves killing an innocent person. But it is hard to see how *Incomparable* would yield the impermissibility of one option and not the other. Either both options would be impermissible, or neither would be. This leads to the second problem with using *Incomparable* to avoid *life for headaches*. Whether *Incomparable* leads to the denial of *life for headaches* depends on which formulation of the consequentialist criterion of permissibility we employ. Consider the following two formulations:

> *Con 1:* An act is permissible iff it is at least as good as all the available alternatives.

> *Con 2:* An act is permissible iff there are no better alternatives.

If we accept *Incomparable*, and consider the situation in which we either kill one innocent person or allow a vast number of headaches to continue, we must still decide which criterion of permissibility to employ. According to *Con 1*, the killing is impermissible, since it is false that it is as good as or better than the only available alternative. Of course, the alternative, allowing the headaches to continue, is also impermissible, for the same reasons. In which case, we have a moral dilemma. On the other hand, according to *Con 2*, the killing is permissible, since there are no better alternatives.

II

A more promising tactic for a consequentialist who wishes to avoid *life for headaches* is to maintain that the loss of an innocent life is worse than a vast number of headaches, than *any* number of headaches, in fact. Consider the claim in this form:

> *Less:* The state of affairs constituted by *any* number of fairly minor headaches is less bad than even one premature death of an innocent person.

A consequentialist who accepts *Less* denies that it is ever permissible to kill an innocent person in order to avoid a number of headaches, no

matter how vast. According to *Less*, if two worlds differ only in that the first contains a vast number of fairly minor headaches not contained in the second, and the second contains one more premature death of an innocent than the first, the second world is worse than the first.

Although *Less* is initially plausible, it is problematic. Roughly, *Less* says that a death is worse than any number of headaches. Consider now a much stronger principle:

> *Less′:* For any misfortune, x, to befall a person, that is worse than that person suffering the pain of a fairly minor headache, the state of affairs constituted by *any* number of fairly minor headaches is less bad than even one person suffering x.

Less′ is wildly implausible. Roughly, it says that anything worse than a headache is worse than any number of headaches. Suppose that a mild ankle sprain is worse than a mild headache. According to *Less′*, it is worse that one person suffer a mild ankle sprain than that billions suffer mild headaches. Clearly, no consequentialist will (or should) accept this. The rejection of *Less′* entails that there are misfortunes that are worse than mild headaches, that nonetheless can be individually outweighed by a sufficient number of mild headaches. This is relatively uncontroversial. A mild ankle sprain is a good candidate for such a misfortune. Likewise, it is pretty clear that there are misfortunes which are worse than mild ankle sprains, but which nonetheless can be individually outweighed by a sufficient number of mild ankle sprains. Perhaps a broken ankle is such a misfortune. Even though it is worse that one person break her ankle than that she mildly sprain it, it is worse that many people have mild ankle sprains than that one has a broken ankle. But this process of escalation can be continued. For each misfortune short of the worst possible one, there is a worse misfortune that can be individually outweighed by a sufficient number of the lesser one. In particular, it seems plausible that there is some misfortune short of death, perhaps some kind of mutilation, that can, if suffered by enough people, outweigh one death. Consider now a sequence of judgments, *S*, that begins as follows: one death is better than n^1 mutilations; n^1 mutilations are better than n^2 xs (where x is some misfortune less bad than mutilation). *S* continues with the first term of each comparison being identical to the second term of the previous comparison, until we reach the last two comparisons: n^{m-2} broken ankles are better than n^{m-1} mild ankle

sprains; n^{m-1} mild ankle sprains are better than n^m mild headaches. If we have *S*, we can conclude, by the transitivity of 'better than' that one death is better than n^m mild headaches. In which case, we must reject *Less*.

How might a consequentialist attempt to salvage *Less*? There are two possibilities that I will discuss. First, she might deny the existence of any true *S*. Perhaps there are some misfortunes so bad that even one such misfortune is worse than any number of any lesser misfortunes. Death is the most likely candidate for such a super misfortune. There is something special about death. As Clint Eastwood says in *Unforgiven*, "It takes away all a man has, and all he's gonna have." Unpleasant as even severe mutilation is, it is still worse that one person die than that any number are mutilated. In which case, the first comparison in *S* will be false. This view of death may appeal to those who believe what students in introductory classes say when they claim that life is 'invaluable' or 'infinitely valuable,' but is it really plausible? Can anyone who really considers the matter seriously, especially a consequentialist, honestly claim to believe that it is worse that one person die than that the entire sentient population of the universe be severely mutilated? My concern here is not to persuade someone who claims that even one death is worse than any number of any lesser misfortunes. I merely wish to explain what is involved in rejecting *S*. Perhaps the break in the sequence of misfortunes could occur at some later point. Perhaps there is some misfortune, short of death, that is worse than any number of any lesser misfortunes. This seems even more implausible, though, than the claim that death is worse than any number of any lesser misfortunes.

III

Even if we accept *S*, we may yet attempt to hang on to *Less*. Recall that the rejection of *Less* followed not from *S* alone, but from *S* and the transitivity of 'better than.' If we reject the transitivity of 'better than,' we may accept both *S* and *Less*. Consider a simpler sequence, *S'*: one death is better than a million broken legs; a million broken legs are better than ten billion mild headaches. If we accept both *S'* and the transitivity of 'better than,' we must accept that one death is better than ten billion mild headaches, and thus reject *Less*. But if we reject the transitivity of 'better than,' we may accept both *S'* and *Less*.

Not only does the claim that 'better than' is intransitive allow a consequentialist to accept both *Less* and the highly plausible *S*, but it also provides a way of accommodating moral dilemmas within a consequentialist framework. Suppose we accept both *S'* and the claim that ten billion mild headaches are better than one death. Consider the following scenario:

> *Intransitive Trolley:* A three-car trolley is speeding out of control toward a three-way fork in the track. The driver, who is a maximizing consequentialist, cannot stop the trolley. If he does nothing, one car will hurtle down each track. He does, however, have three other options, corresponding to three buttons on his control panel. If he pushes button 'L,' all three cars will go down the left track. If he pushes 'M,' all three cars will go down the middle track. If he pushes 'R,' all three cars will go down the right track. If one or more cars enter the left fork in the next minute, exactly one innocent person will be killed. If one or more cars enter the middle fork in the next minute, exactly one million innocent people will suffer broken legs. If one or more cars enter the right fork in the next minute, exactly ten billion innocent people will suffer mild headaches.

What should the driver do? One of his options is clearly worse than all the others. Doing nothing is worse than pushing any of the buttons. However, there is no best option, no option that is at least as good as all other options, and no option than which there are none better. Of the three button-pushing options, each one is better than one other and worse than one other. L is better than M, but worse than R. M is better than R, but worse than L. R is better than L, but worse than M. There is no permissible option. The driver is thus faced with a moral dilemma. Notice that it does not matter which version of the permissibility requirement we employ, *Con 1* or *Con 2*. In this sense, the denial of the transitivity of 'better than' is a more effective way of accommodating moral dilemmas within a consequentialist framework than the claim that some values (or options) are incomparable.

Should a consequentialist embrace intransitivity? The considerations of the previous paragraph will probably not sway many consequentialists, who tend to reject the claim that there are moral dilemmas[5] (as opposed to moral conflicts, in which even the best option is regrettable).

5. See, for example, Alastair Norcross, "Should Utilitarianism Accommodate Moral Dilemmas?" *Philosophical Studies* 79, no. 1 (July 1995): 59–85.

Is *Less* such a compelling principle that we should reject something as overwhelmingly plausible as the transitivity of 'better than' in order to accommodate it? There are, in fact, considerations that directly undermine *Less*. For example, it is reasonable to suppose that, at least sometimes, raising the speed limit on highways is better than keeping it at its current level. The benefits of increased speed limits, however, are merely increased convenience for many, while the losses are the deaths of a few. I will postpone a fuller consideration of this example until later in this article. I mention it here simply to raise the possibility that the attractiveness of *Less* may be no more than skin deep. The transitivity of 'better than,' on the other hand, is deeply entrenched. In fact, some may even regard it as an analytic truth. Clearly, then, additional argument is needed to convince a consequentialist to embrace intransitivity.

IV

In this and the next section, I will consider two kinds of attempt to argue for the intransitivity of 'better than.' First I will examine a sorites-type argument, presented by Warren Quinn, that uses a sequence of imperceptible differences in utility that add up to a vast change. Second, I will examine an argument of Larry Temkin's, originally devised by Stuart Rachels, that our experiences of pain form a continuum, such that a sequence of definite changes for the worse can add up to a change for the better.

Warren Quinn, in "The Puzzle of the Self-Torturer,"[6] presents the example of a self-torturer hooked up to a device "that enables doctors to apply electric current to the body in increments so tiny that the patient cannot feel them. The device has 1001 settings: 0 (off) and 1 . . . 1000." At each setting, the self-torturer is offered $10,000 to advance to the next one. Since he can feel no difference between adjacent settings, he always prefers to advance a setting, and receive the money, rather than to stay put. However, by the time he reaches 1000, he is in so much pain that he prefers to return to 0 and relinquish all the money.

Quinn uses the example to argue, not that 'better than' is intransitive, but that the torturer's stepwise preferences are intransitive. However, if Quinn is correct in his claim that the torturer really is just as comfortable at each setting as at the previous one, his example provides a conse-

6. *Philosophical Studies* 59 (1990): 79–90.

quentialist, such as a utilitarian, with a reason to reject the transitivity of 'better than.' Consider any choice between adjacent settings, S and S+1. If the torturer chooses S+1, he will not be any less comfortable than if he stays at S. Furthermore, he will get something he cares about, $10,000, that he will not get if he stays at S. There are no respects in which he is worse off at S+1 than at S, and there is at least one respect in which he is better off. The self-torturer will, therefore, be better off all-things-considered at S+1 than at S. If we consider two worlds, W and W+1, that differ only in that the self-torturer is at S in W and at S+1 in W+1, we should judge that W+1 is better than W. However, according to Quinn, the self-torturer is worse off at 1000 than at 0. So the world in which he is at 1000 is worse than the one that differs from it only in that he is at 0. Thus, intransitivity looms.

The argument of the previous paragraph depends crucially on the claim that the torturer is just as comfortable at each step as at the previous one. Why should we accept this? Quinn considers several reasons to reject this, and dismisses them, rather too hastily. For the sake of brevity, I will consider just two of them. We might point out that Quinn's claim is only plausible while we focus on the torturer's introspective ability to distinguish between his comfort level at adjacent settings. There may, however, be behavioral evidence, unavailable to introspection, that the torturer's comfort level is declining— "he might look less comfortable, or be grouchier." Quinn's response to this possibility is simply to stipulate that he is interested in a case in which "the individual increments of current are too small even to have these effects." This hardly gives the objection a run for its money. If we consider all the behavioral evidence there might be, it is not clear how the increments of current could be so small that no increase made any difference over the previous setting. At 0 the torturer is not screaming, even a little bit. At 1000 he is screaming a lot. If his behavior is no different at 1 than at 0, he is not screaming, even a little bit, at 1. Likewise, he is not screaming, even a little bit, at 2. And so on. So he is not screaming, even a little bit, at 1000. But he is. Clearly, there must be objective differences that are made by some of the increments of current. To say that there is an objective difference is not necessarily to say that the torturer must move from determinately not screaming to determinately screaming. Perhaps at one setting he is determinately not screaming. At the next, it is no longer determinately true that he is not screaming. This would still con-

stitute an objective difference. If so, it is possible that his comfort level really does decline when he moves from one setting to the next, no matter how small the increment. In which case we do not have to accept that he is really better off at each setting than at the previous one.

Quinn also considers the following criticism:

> *We are ignoring the effects of "triangulation":* The self-torturer can triangulate a difference between s and s+1 if he can find some third setting s′ that feels the same as s but better than s+1. And if he can use o to triangulate such a difference, then it is obvious that his comfort, at least compared to o, declines in stepping from s to s+1.

His response is to stipulate, again, that he is interested in a case in which the individual increments of current are "too small to make *any* difference in comfort, even one that can be detected only by triangulation." This is puzzling. After all, no matter how small the increments of current, the self-torturer eventually reaches some settings that feel worse than o. At the very least, o feels better than the last setting. It is also stipulated that o feels the same as 1. If he simply considers each setting in turn, why will there not be a first one that feels worse than o, or at least a first setting of which it is no longer determinately true that it feels the same as o? Quinn responds as follows:

> But, empirically speaking, his preferences as between *s* and o can exhibit various kinds of *indeterminacy.* Not only is there no empirically determinable *first* setting that he disprefers to o, there is no empirically determinable *first* setting at which these preferences become indeterminate.

Imagine the self-torturer comparing successive settings with o. In front of him is a control panel with three buttons, labeled 'Yes,' 'No,' and 'It's Indeterminate.' After performing each comparison, he is asked, "Does this setting feel the same as o? Push the appropriate button within the next sixty seconds." If he responds within one minute, the message corresponding to the button pushed appears on a screen. Otherwise no message appears. If this procedure is performed, there will clearly be a first setting at which 'yes' does not appear on the screen.[7]

7. For similar, steplike, arguments against vague objects, see Mark Heller, "Against Metphysical Vagueness," *Philosophical Perspectives* 10, *Metaphysics* (1996), and Roy Sorensen, "A Thousand Clones," *Mind* 103 (1994).

But perhaps I am being unfair to the case. I have only given the self-torturer three options (four, if you include not pressing any button). Why assume that the only possibilities for any setting are that it feels the same as 0, it does not feel the same as 0, and it is indeterminate whether it feels the same as 0? Perhaps there are settings for which it is indeterminate whether it is indeterminate whether it feels the same as 0. Call such a state of affairs '2indeterminacy.' But then, will there not be a first setting of which it is 2indeterminate whether it feels the same as 0? Not necessarily. Perhaps there are settings of which it is indeterminate whether it is 2indeterminate whether it feels the same as 0. Call such a state of affairs '3indeterminacy.' This suggests that any claim to the effect that there must be a first setting of which it is *n*indeterminate whether it feels the same as 0 can be countered with *n*+1indeterminacy. This move will work for any *n*.

However, now let us say that it is *super*indeterminate whether a setting feels the same as 0, just in case there is *any* indeterminacy *of any kind* or *at any level* about whether the setting feels the same as 0. Can we now block the claim that there will be a first setting of which it is *super*indeterminate whether it feels the same as 0? Can there be settings of which it is indeterminate whether it is *super*indeterminate whether they feel the same as 0? Clearly not. The postulation of indeterminacy about *super*indeterminacy is self-defeating. So now it appears that there will be a first setting, of which it is either *super*indeterminate whether it feels the same as 0, or false that it feels the same 0. In moving from the previous setting to this setting, there is an objective change in the torturer's comfort.

I conclude that Quinn has not succeeded in showing that 'just as comfortable' is intransitive, and thus that he has not provided a consequentialist with a reason to reject the transitivity of 'better than.'

V

Larry Temkin also argues that 'better than' is intransitive; his argument centers on a variation of an example of Stuart Rachels,[8] and rests on the following three claims.

8. Larry S. Temkin, "A Continuum Argument for Intransitivity," *Philosophy & Public Affairs* 25, no. 3 (Summer 1996): 175–210. All page references in the text are to this article, unless otherwise stated. Temkin offers a different argument for intransitivity, based on

> Claim 1: for any unpleasant or "negative" experience, no matter what the intensity and duration of that experience, it would be better to have that experience than one that was only a little less intense but twice as long.
> Claim 2: there is a continuum of unpleasant or "negative" experiences ranging in intensity, for example, from extreme forms of torture to the mild discomfort of a hangnail.
>
> Claim 3: a mild discomfort for the duration of one's life would be preferable to two years of excruciating torture, no matter the length of one's life. (179)

Claims 1–3 all "depend on an appropriate 'other things equal' clause" (182). Now, we are to consider a sequence of comparisons of lives, all overall worth living, A with B, B with C, C with D .... X with Y, that begins as follows: "First, compare two lives A and B. Suppose that A and B are both lengthy—perhaps, indeed, *very* lengthy—and that A and B are similar, except that A contains two years of excruciating torture, B four years of torture whose intensity is almost, but not quite, as bad as A's.... Next, compare B with C, where C stands to B as B stands to A" (180). The sequence continues, with each successive life containing a pain that is a little less intense but twice as long as the pain in the previous life. At the end of the sequence, "one would be comparing two alternatives, say, X and Y, such that X had an annoying hangnail for a *very* long time—perhaps thousands of years[9]—and Y had a hangnail that was almost, though not quite, as unpleasant as X's, but that lasted twice as long" (180). In accordance with claim 1, A is better than B, B is better than C . . . , X is better than Y. If 'better than' is transitive, A is better than Y. But claim 3, and the intuitions of many when first presented with the case, tell us that A is not better than Y. Thus transitivity is threatened.

---

what he calls the '*Person-Affecting Principle*,' in "Intransitivity and the Mere Addition Paradox," *Philosophy & Public Affairs* 16, no. 2 (Spring 1987): 138–87. I discuss this argument in my "Intransitivity and the Person-Affecting Principle," American Philosophical Association annual meeting, Central Division in Pittsburgh, May 1997. Rachels's version of the example appears in "Reconceiving 'Better Than,'" American Philosophical Association annual meeting, Pacific Division in San Francisco, April 1995, and "Counterexamples to the Transitivity of 'Better Than,'" *Australasian Journal of Philosophy*, forthcoming.

9. Actually, if the sequence contains lives corresponding to all the letters A through Y, X's hangnail would last for more than 16 million years, and Y's for more than 32 million years.

This argument is provocative, and initially compelling. However, I remain unconvinced. Claims 1 and 2 are relatively unproblematic, so long as they are interpreted loosely with the relevant 'other things being equal' clauses.[10] It is claim 3 that I find problematic. First, let us be clear about what it says in the context of this example. Claim 3, as it applies to the comparison of A with Y, tells us that a life that is 32 million years long, overall worth living, and that contains the discomfort of a mild hangnail for the entire 32 million years is preferable to a life that is also overall worth living, and is just like it,[11] except that, for two of the 32 million years it contains excruciating torture instead of hangnail pain, and for 31,999,998 years it contains neither torture nor hangnail pain.

Both claim 3 and the judgment that A is worse than Y are initially intuitively appealing. But how trustworthy are our intuitions, when it comes to a comparison between A and Y? How many of us have had the misfortune to experience excruciating torture at all, let alone for two years? I have experienced pains no more severe than a broken wrist, torn ankle ligaments, or an abscess in a tooth. These were pretty bad, but I have no doubt that a skilled torturer could make me experience pains many times as bad. I have experienced a mildly annoying hangnail. When I try to imagine two years of pain that is many times worse than the worst pain I have ever experienced, and I consider whether any amount of hangnail pain could outweigh it, the immediate answer is no. But just what am I considering? Two years is a fairly sizable portion of a human life. How could any life with that large a portion spent in unimaginable pain be better than a life with only a mildly annoying hang-

10. Thus, we would not insist on an infinitely divisible continuum for claim 2, or on the less intense experience of claim 1 being twice as long, as opposed to even longer. Temkin makes these points himself:

> More to the point, for any imaginable unpleasant experience, we can imagine another that would be noticeably better, but still sufficiently bad that we would prefer the former to twice (or three or five times) as much of the latter. This suggests that even if, phenomenologically, there isn't a smooth continuum, there is still a discontinuous spectrum of possible experiences ranging from A to Y that is sufficient for my purposes. (183–84)

11. The two years of torture have no effect on the rest of A. There are no painful memories afterwards, and no dreaded anticipation beforehand. Likewise, the hangnail pain remains constant throughout Y. It does not "eventually have the effect of the so-called Chinese water torture—where a steady drip of water ultimately drives one crazy" (181). But neither does one become used to it, so that eventually an extra day, or year, or even million years makes no difference. If that were the case, it would not be clear that Y was worse than X.

nail throughout? But in comparing A with Y we are supposed to imagine lives many times longer than any person has ever lived. I suggest that our intuitions about the comparison of A with Y are unreliable in at least two respects. First, it could be that our intuitions overestimate the significance of the two years of torture in A, because we simply cannot imagine what it would be like to live for 32 million years. Two years may be a sizable portion of a standard human life, but it is a tiny fraction of 32 million years. Proportionally, it is the equivalent of 2.6 minutes of torture in an eighty-year human lifetime. (Would A seem worse than Y, if the torture were to be experienced for ten seconds every five years? Clearly not.) Second, we underestimate the significance of the 32 million years of hangnail pain in Y. I have trouble imagining what it would be like to have a hangnail for even a few years, let alone an entire ordinary lifetime. And yet, when I try to compare A with Y, I am supposed to imagine having a hangnail continuously for 32 million years!

Temkin considers at least the second criticism of the previous paragraph. His reply, which applies equally to the first criticism, is twofold. First, he claims that, even if claim 3 is false, his argument will still work, because A and Y will not be anything like as lengthy as I suggest. Second, he defends claim 3 directly. I will examine his first reply first.

Temkin claims that we may be able to traverse the pain spectrum from excruciating torture to mildly annoying hangnail in a relatively small number of steps, each of which is small enough to fit the original structure of the argument. Perhaps, he suggests, our experience of pain is analogous to our experience of color:

> Consider the following spectrum: red, reddish orange, orange, orangish yellow, yellow, yellowish green, green, greenish blue, blue, indigo-blue, indigo, indigo-violet, violet. It is hard to deny that, phenomenologically, the gaps between the adjacent members of this spectrum are pretty small, though still clearly perceptible. Yet it would only take twelve steps to get from one extreme of the spectrum to the other. (186)

If we could get from excruciating torture to mildly annoying hangnail in twelve steps that are small enough to fit the structure of the original example, A and Y (or whatever the last life in the sequence would be called) would be a 'mere' 8192 years long. That is certainly less of a stretch for our human imaginations than 32 million years, but I am still

not ready to trust my (or Temkin's) intuitions on the comparison of A with Y. But, let us grant, for the sake of argument, that we can confidently proclaim that 8192 years of a hangnail is better than two years of torture and 8190 years of neither torture nor hangnail (other things being equal). Why should Temkin's color comparison persuade us that A and Y could be as short as 8192 years? Temkin concedes that the color analogy is "purely suggestive," but he adds that he "would be surprised if our phenomenological powers of discrimination differed radically across our sense modalities" (188).

But what would it be for the case of pain to be analogous to that of color, in the way required for Temkin's argument? Say we divide the distance from torture to hangnail into twelve equal steps.[12] Likewise, we have twelve steps from red to violet.[13] Now, the crucial question is how small are the steps across the pain spectrum? Well, we say, if they are analogous to the steps across the color spectrum, they are pretty small. After all, the step from, say, yellow to yellowish green is pretty small. But, this is too vague to be any help. We need to know whether the steps across the pain spectrum are *small enough for the purposes of Temkin's argument*. That is, it must be clear that x years of one pain is preferable to 2x years of the pain that is just one step down from it. How do we pursue the analogy with color? Is one year of yellow clearly better than two years of yellowish green? Or perhaps we should ask whether one year of yellowish green is clearly better than two years of yellow. Neither question makes sense. This is precisely the problem with arguing, even suggestively, from the case of color to the case of pain. Comparisons of pains have natural evaluative import. If one pain is greater than another, it is, other things being equal, preferable to experience the latter. The pain spectrum has a natural evaluative direction, from greater pains to lesser pains. There is nothing analogous with color comparisons or the color spectrum. There is no natural sense in which red is preferable to or greater than orange, or vice versa.

Perhaps we could try to force evaluative significance on the color spectrum with a thought experiment. Imagine beings whose qualitative

12. The steps must be equal in the proportionate sense. In the sense in which the step from 5 to 10 is equal to the step from 10 to 20, not to the step from 10 to 15. Call the sense in which the step from 5 to 10 is equal to the step from 10 to 15 additive equality.

13. Here we have the first problem with the analogy. It is not clear what, in the color case, would correspond to the difference between proportionate and additive equality in the pain case. If anything, Temkin's twelve steps across the color spectrum seem to be equal in the additive sense, but he needs them to be equal in the proportionate sense.

states are just like ours, with the following addition. The experience of color is painful for them. The pain of experiencing red is the same intensity as excruciating torture. The intensity diminishes through the color spectrum, with the pain of experiencing violet having only the intensity of a mildly annoying hangnail. The qualitative feel of a color experience is the same as ours, with the addition of the pain. Imagine now the difference between the experience, for such a being, of yellow and of yellowish green. Since part of those experiences is the same as ours, the nonpainful part that is, we have a perfectly clear idea of how similar they are in that respect. But the respect we are interested in is how much better is the experience of yellowish green than the experience of yellow. In particular, how long a yellowish green experience would it take to outweigh, say, one hour of yellow experience. To answer that question, we need to know how the pain parts of the experiences compare. So we need to divide the pain spectrum, from torture to hangnail, into twelve equal (proportionate) steps, and ask how much of one pain would it take to outweigh, say, one hour of the next pain up. But this is the question the color analogy was supposed to help us with. Instead, it seems that we cannnot make sense of the color analogy without first answering this question.

I have been arguing that Temkin's color analogy gives us no reason to believe that we could traverse the pain spectrum in as few as twelve steps, all of which are small enough to keep lives A and Y to (roughly) imaginable lengths.[14] However, perhaps we need not settle the question of how many (small enough) steps it would take for *us* to get from excruciating pain to a mildly annoying hangnail. Perhaps, as a matter of empirical fact, it would take us so many steps to get from one end of the

14. Temkin also tries a variation of his example, that envisages a cardinal ranking of pains from 1 (mildly annoying hangnail) to 100 (excruciating torture). On this scale, a moderately uncomfortable limp is 11. If we start with two days of excruciating torture, and move in steps that decrease the intensity of the pain by 20 percent but double its duration, we reach 5.6 years of a moderately uncomfortable limp in ten steps. Temkin claims that, other things being equal (the torture leaves no memories, the limp has no side-effects, etc.), two days of torture is worse than 5.6 years of the limp. Even if I agreed with that judgment, I would find the argument unconvincing. On Temkin's scale, excruciating torture is fractionally more than nine times worse than a moderately uncomfortable limp. I have had a moderately uncomfortable limp, when I was recovering from torn ankle ligaments. In my judgment, the pain of the freshly torn ligaments was *far* more than nine times as bad as the limp. Furthermore, as I stated above, I am quite sure that an experienced torturer could make me experience pain *many* times worse than the pain of freshly torn ligaments.

pain spectrum to the other that A and Y would be millions (or even billions) of years long. Temkin's response is that, even if this is true, it is surely only contingently true. All he needs is the bare *possibility* of beings whose experience of pain is different enough that A and Y, for them, would be only hundreds, or perhaps thousands, of years long.[15] Imagine beings whose experience of pain is sufficient for Temkin's purposes. Call them 'intransitives.' If it is metaphysically possible for there to be intransitives, the axiom of transitivity for 'all things considered better than' is false. Even if we are reasonably sure that *we* are not intransitives, surely it is (metaphysically) *possible* for there to be intransitives. So, the axiom of transitivity is false.

This argument is simple and appealing, but we must be careful not to slide from epistemic possibility to metaphysical possibility. Suppose, for the sake of argument, that we just do not know whether there are intransitives. For all we know, there are intransitives. That is, it is epistemically possible that there are intransitives. Or perhaps we are fairly sure that there are no intransitives, but, for all we know, there may be. So, it is epistemically possible that there may be intransitives. Nothing follows from either claim of *epistemic* possibility about the *metaphysical* possibility of intransitives. (Imagine an agnostic, who holds that, for all he knows, God exists, and, for all he knows, God does not exist, and who believes that God, if there is one, is a necessary being. If he slides from the epistemic to the metaphysical possibility of God's existence and nonexistence, he will be committed both to the view that God exists and the view that God does not exist!)

Why should we believe that it is metaphysically possible for there to be intransitives? Here is a bad argument, which might underlie an intuition that intransitives are metaphysically possible. It is clearly a contingent matter that we experience pain the way we do. We could have been constructed differently, and have had different experiences of the pain spectrum. So, we could have had the experiences of the intransitives. But, even if it is true that our experiences of pain could have been different, it does not follow that they could have been different in just *any* way. The main quad of Christ Church, Oxford, could have had a differ-

15. "Unless it is metaphysically impossible for the pain and color modalities to be analogous in the way suggested, there are some possible creatures for whom my argument would work—even if it does not, in fact, work for us. This is all one needs to undermine the claim that 'all things considered better than' *must* be transitive" (188).

ent shape and/or dimensions. But here is one combination of shape and dimensions it could not have had: square, with sides of 200 feet each and a diagonal of 100 feet.

Here's another bad argument that, until exposed, could be driving our intuitions. Even if our pain spectrum spans such a large range from excruciating torture to mildly annoying hangnail, we might have been constructed differently, so that the range from the worst pain we could experience to the least pain was much smaller. In particular, we might have been able to get from the worst pain to the least pain in a small number of steps, say, twelve, each of which was small enough that a certain duration of one pain was clearly preferable to twice the duration of the next pain down. Call this possible feature of pain experience '*Q*.' Since having *Q* is all it takes to be an intransitive, there could have been intransitives. The problem is that *Q* is not all it takes to be an intransitive. Our experience of pain needs to have *Q* and the following further feature, call it *'I'*: 8192 years of the least pain is clearly better than two years of the worst pain and 8190 years of neither the worst pain nor the least pain (other things being equal). *I* would seem to require a pain spectrum as broad (roughly) as ours. Many possible creatures could have *Q* but lack *I*. Consider a creature whose pain spectrum ranges from the pain of a broken leg to the pain of a sprained ankle. That is, it ranges from the pain that a typical human suffers from a broken leg to the pain that a typical human suffers from a sprained ankle. Such a creature would probably have *Q*. They could get from their worst pain—the pain of a broken leg—to their least pain—the pain of a sprained ankle—in twelve small enough steps. But would 8192 years of sprained ankle pain be clearly better than two years of broken leg pain and 8190 years of neither sprained ankle pain nor broken leg pain (other things being equal)? Clearly not. Temkin's argument depends on the ends of the pain spectrum being roughly as far apart as excruciating torture and a mildly annoying hangnail. While it may be fairly easy to imagine beings whose pain spectrum is narrow enough that they could get from one end to the other in twelve small enough steps, it is not so easy to imagine beings who could span a spectrum that is roughly as broad as ours in twelve such steps.

Let me illustrate what I take to be the problem with the claim that intransitives are metaphysically possible. My back yard is about 120 feet long. My stride is about 3 feet. So I can traverse the length of my back

yard in 40 strides. Could it have been the case that some creature could have traversed my back yard in 10 strides? Yes. Perhaps there could have been a creature with a 12-foot stride, or perhaps my back yard could have been only 30 feet long. But, could there have been a creature with the same length stride that I actually have, who could have traversed a backyard of the same length as my actual back yard in only 10 strides? Clearly not. When we ask whether there could be intransitives, we are asking whether there could be creatures, some of whose experiences violate the axiom of transitivity for 'all things considered better than.' If we believe, as I do, that 'all things considered better than' is transitive for all possible alternatives, we have a strong reason to reject the metaphysical possibility of intransitives. If we reject the axiom of transitivity for 'all things considered better than,' we lack at least that reason for rejecting the metaphysical possibility of intransitives. I do not see how we can consider the metaphysical possibility of intransitives, in isolation from our belief in the axiom of transitivity, in an attempt to shape that belief.

I conclude that we have no good reason to suppose that any possible pain experiences could span the pain spectrum from torture to hangnail in few enough small enough steps that A and Y were 'merely' a few hundred (or even thousand) years long. However, it is still possible to accept Temkin's argument, if we are prepared to accept claim 3, despite the problem of trusting our intuitions regarding pains of unimaginably long duration. So, I turn now to a direct consideration of claim 3.

Perhaps we think that two years of excruciating torture is so bad that only more of the same kind of suffering could be worse than it. And we think this, even if the two years of torture is a tiny fraction of an enormously long life, the rest of which is happy enough (perhaps blissfully happy) to make the life overall worth living. B is worse than A, because B contains the same kind of suffering as A, albeit a little less intense, and lots more of it. The suffering in Y, on the other hand, even though almost unimaginably lengthier than the suffering in A, is clearly not of the same kind. So Y cannot be worse than A.[16] A little thought will tell us that this line of reasoning will not justify the claim that A is worse than Y. Recall that A is overall worth living. It is, therefore, better than any life that is not overall worth living. Is it not possible to alter Y in such

16. Temkin seems to endorse this explanation on p. 194.

a way that it is not overall worth living, without introducing the same kind of suffering as the torture in A? Clearly, a life of unmitigated excruciating torture would not be worth living. But there may be many lives that are not worth living, that contain nothing like excruciating torture at any point. Imagine 32 million years (or even just fifty or so years) of tasteless food, unfulfilling relationships, ill-fitting, drab, uncomfortable clothes, a tedious job with contemptuous and contemptible coworkers, a drafty, dirty one-room apartment with a lumpy bed, a thoroughly uninquiring intellect, and a continuous mildly annoying hangnail. Call such a life "YME." It seems highly likely that YME would not be worth living, and so would be worse than A. (If you think that YME would be worth living, just keep adding details until you get to a life you consider not worth living. I think you will find you can do it, without introducing anything like excruciating torture.) Yet YME contains none of the same kind of suffering as A. So, two years of excruciating torture *can* be outweighed by enough fairly minor suffering and deprivation. Why, then, can it not be outweighed by an enormously lengthy hangnail pain? My example, of course, does not prove that two years of excruciating torture can be outweighed by some amount of hangnail pain, but it does undermine a principle which may well be the basis for our intuitive assent both to claim 3 and to the specific judgment that A is worse than Y.[17]

Let us look now at Temkin's explanation of why two years of torture cannot be outweighed by any amount of hangnail pain:

> [I]n essence, I think significant amounts of torture have lexical priority over any amount of a hangnail, just as I think lives have lexical priority over lollipop licks. My model for this is something like the following. Torture's badness might range from 0 to 10, depending on its duration, with two years of torture being, say, a 7. A hangnail's badness might range from 0 to 1. Prolonging a hangnail increases the value of the decimal places representing its "badness score," but the fundamental gap between 1 and 7 is never affected. (191)

What do the numbers signify in this account? The claim that two years of torture is a 7, on a scale of 0 to 10, does not tell us that any life contain-

17. Perhaps Temkin could respond to my example of YME by altering his example, so that all the lives A . . . Y are not worth living. Rachels's version of the example (the most recent version that I have seen) has this structure. Such a move has other difficulties, though, that I do not have the space to explore here.

ing two years of torture is equally bad, nor does it tell us that four years of torture is not, other things being equal, twice as bad as two years of torture. It tells us that any life containing two years of torture is worse than any life of equal length, which is just like it, except that instead of torture it contains some amount of pain whose badness ranges from 0 to less than 7. Since Y is of equal length with A, and just like it except that, instead of two years of torture, Y contains a certain (very large) amount of a pain, whose badness ranges from 0 to 1, A must be worse than Y.

Presumably, there are some pains worse than a hangnail whose badness ranges from 0 to less than 7. A will be worse than any life in the sequence that contains such a pain. X, for example, contains about 16 million years of annoying (as opposed to mildly annoying) hangnail pain. I should think Temkin would want to say that the badness of annoying hangnail pain ranges from 0 to considerably less than 7. Thus, A must be worse than X. There must also be some pains less intense than excruciating torture, whose badness ranges from 0 to more than 7. The slightly less excruciating torture in B would be such a pain. Since we know that B is worse than A, four years of the pain in B must be more than 7, even if not by much. Let us say, then, that four years of the pain in B has a badness score of 7+x. Now, since we also know that C is worse than B, eight years of the pain in C must have a badness score of more than 7+x, even if not by much. Let us say, then, that eight years of the pain in C has a badness score of 7+x+y. The problem now is obvious. How are we ever going to get to a pain with a badness score of less than 7? Since each life in the sequence is worse than the previous one, the badness score for the specific duration and type of pain in each life must be higher than that for the corresponding pain in the previous life. Thus, either the range for a mildly annoying hangnail is from 0 to more than 7, which contradicts this model for understanding claim 3, or at least one life in the sequence is clearly better than the previous one, which contradicts claim 1.

Perhaps Temkin's numerical model for explicating claim 3 is merely infelicitous. Let us try to express the same idea without recourse to numbers. A is worse than Y, because any life containing two years of excruciating torture is worse than any life of equal length that is just like it, except that, instead of the torture, it contains some (perhaps very lengthy) duration of mildly annoying hangnail pain. Let us express this

as the claim that A has the property of being worse than any amount of hangnail pain.[18] In order for this to be genuinely explanatory, we are to think of this as an intrinsic feature of A, which has relational consequences. It is like the claim that Tom is over six feet tall, rather than the claim that Tom is taller than Dick, who happens to be five feet tall, and Harry, who happens to be four feet tall.[19] Call all and only lives that have the property of being worse than any amount of hangnail pain 'miserable.' Is B miserable? I should imagine that Temkin would claim that it is. X, on the other hand, clearly is not miserable. It cannot be, since, by claim 1, X is better than Y. Consider now the following claim:

> Claim 4: between any two lives, x and y, of equal length, that differ only in that x has a certain duration of one pain and y has a certain duration of a pain of different intensity, if x is miserable, then x is better than y only if y is also miserable.

Claim 4 is, on inspection, just as compelling as claims 1–3. (In fact, most would assent to a stronger claim, according to which y is better than x, if y is not miserable.) But now we have trouble reconciling claims 1, 3, and 4. Consider the sequence A . . . Y. According to claim 3, A is miserable. Clearly Y is not miserable. So, there must be a first life, α, in the sequence that is not miserable. By claim 4, then, the previous life in the sequence is not better than α. But this contradicts claim 1. Perhaps indeterminacy will rear its ugly head again. Maybe there are no adjacent lives in the sequence such that the first determinately is miserable, and the second determinately is not. However, there will be a first life, β, of which it is no longer determinately true that it is miserable. According to claim 4, then, it is not determinately true that the previous life in the sequence is better than β. This still contradicts claim 1. Perhaps Temkin could come up with some variations on claims 1 and 3 that are compel-

18. A life is worse than any amount of hangnail pain, just in case it contains some duration of some pain such that, any life containing that duration of that pain is worse than any life of equal length that is just like it, except that, instead of that pain, it contains some duration of mildly annoying hangnail pain.

19. Suppose that having the property of being worse than any amount of hangnail pain were a purely relational feature of A. In that case, the explanation that A is worse than Y because A is worse than any amount of hangnail pain has the same explanatory force as the claim that A is worse than Y because A is worse than X or Y or Z. Another reason to construe the property of being worse than any amount of hangnail pain as intrinsic, is so that claim 4, which employs this property, does not have the appearance of presupposing transitivity.

ling, can coexist happily with claim 4, and would threaten transitivity. However, I suspect that I could come up with variations on claim 4 that would once again threaten the harmony.

The previous two arguments share a common structure. If Temkin is to explain why A is worse than Y, rather than simply assert that it is, he must appeal to a property possessed by A but not by Y (e.g., having a badness score of more than 1, being miserable). Given the 'other things being equal' clause that applies to all the lives in the sequence, the possession by A but not by Y of the property is sufficient for the judgment that A is worse than Y. There will, however, be a first life in the sequence that does not possess the property. Therefore, there will be an adjacent pair of lives, say m and n, such that m does but n does not possess the property. Therefore, n is not worse than m. But this contradicts claim 1.

I will close this section with a direct challenge to claim 3. First, let me repeat the claim.

> Claim 3: a mild discomfort for the duration of one's life would be preferable to two years of excruciating torture, no matter the length of one's life.

Imagine a life that is 500 years long and overall worth living, but that contains two years of excruciating torture. (If you think that no life of a mere 500 years which contained two years of torture could be worth living, simply make it longer and adjust the rest of the numbers in this example accordingly.) The 498 years not spent in torture would be tolerably good, certainly good enough to make the life overall worth living. Call the condition during the 498 years 'bliss.' A life of two years of torture and 498 years of bliss is, therefore, worth living. What if we add 500 more years of bliss? Would such a life—2 years of torture and 998 years of bliss—not be better than the original life? Perhaps such a life, call it 'A,' would even be well worth living. According to claim 3, though, 1000 years of hangnail pain would be better than A. (That is, 1000 years, all of which was just like bliss with the addition of a mildly annoying hangnail.) Claim 3, with the appropriate 'other things being equal' clause, does not, however, entail that *all* lives containing just hangnail pain are better than A. What about a life that consisted entirely of one hour of hangnail pain? Would it be better to live for only one hour, during the whole of which you had a hangnail, or to live for 1000 years, two years

in torture and 998 years in bliss? Given that A is worth living, it seems overwhelmingly likely that some lives containing a hangnail throughout, call them 'hangnail lives,' are not better than A. In particular, it seems likely that there is some hangnail life, call it 'HL,' such that it is rational to be indifferent between living A and living HL.

Suppose that HL is 100 years. (If you think it would have to be much shorter, simply adjust the numbers accordingly.) Now imagine Bart and Lisa choosing in advance how their lives will go, from among a limited set of options. Each is given the choice only between A and HL.[20] Bart chooses HL and Lisa chooses A. Since it is rational to be indifferent between HL and A, neither lives a better life than the other. Suppose now that they are both able to extend their lives in one of two different ways. At the end of the initial segments of their lives, they are both given the choice between living for a certain duration in bliss or living for a longer duration with a hangnail. Living with a hangnail is worse, albeit not by much, than living in bliss. Therefore, for any particular duration of bliss, t, there is a longer duration, t+, such that it is rational to be indifferent between living for t in bliss and living for t+ with a hangnail. The increased length of t+ over t compensates for the slight decrease in quality between the blissful life and the hangnail life (recall that the hangnail life is worth living). In particular, there is a length of bliss, such that it is rational to be indifferent between that length of bliss and that length plus 900 years of a hangnail. Suppose that length to be 10,000 years (again, nothing rests on the particular choice of numbers). It is, therefore, rational to be indifferent between 10,000 years of bliss, call it 'B,' and 10,900 years of a hangnail, call it 'H.'

Bart and Lisa are each offered the choice between B and H. Not only are they both rationally indifferent between B and H, but the experience of B *for Lisa* is just like the experience of B *for Bart*, and likewise for H.[21] Lisa chooses B and Bart chooses H. Lisa's total life now consists of A followed by B, Bart's of HL followed by H. So, Lisa's total life consists of

20. Furthermore, A for Bart is just like A for Lisa, and likewise with HL. For more details on this specification, see the following note.

21. This specification is important in order to construct a pair of lives to which claim 3 is supposed to apply. Bart's hangnail life is supposed to be just like Lisa's bliss, with the addition of the hangnail pain. Since the choice between B and HL is the choice between 10,000 years of bliss and 10,900 years just like bliss with the addition of hangnail pain, if Bart's experience of bliss would not have been like Lisa's, his hangnail experience would not have been just like her bliss with the addition of hangnail pain.

two years of torture and 10,998 years of bliss. Bart's is just like it, except that, instead of the torture, he has 11,000 years of hangnail pain. Given that neither A nor HL was preferable to the other, and that neither B nor H was preferable to the other (and that B, for Lisa, was the same as it would have been for Bart), neither Lisa's nor Bart's total life is preferable to the other. But claim 3 says that Lisa's life is worse than Bart's. So we should reject claim 3.[22]

Given the overwhelming plausibility of the claim that 'all things considered better than' is transitive, and the serious doubts I have raised about Temkin's argument, I conclude that he has provided an interesting and challenging, but ultimately unconvincing argument for intransitivity.

VI

We have seen that neither Quinn's sorites argument nor Rachels's and Temkin's continuum argument can provide good reasons to reject the transitivity of 'better than.' Unless some other argument surfaces, a consequentialist who wishes to embrace the overwhelmingly plausible *S* must reject *Less:*

> *Less:* The state of affairs constituted by *any* number of fairly minor headaches is less bad than even one premature death of an innocent person.

But, for a consequentialist, the rejection of *Less*, coupled with the rejection of incomparability, leads to *life for headaches*:

22. It might be objected that we can construct a pair of clearly unequal composite lives out of choices just like the ones in my example. Suppose that Bart and Lisa are each indifferent between studying the violin and studying the clarinet in college. Bart chooses the violin and Lisa chooses the clarinet. At the end of college, they are both presented with the choice between continuing their studies at Conservatory A or Conservatory B. They are both (rationally) indifferent between the two conservatories. Bart chooses A and Lisa chooses B. It does not follow that Bart's experience of studying the violin in college and in Conservatory A is neither better nor worse than Lisa's of studying the clarinet in college and in Conservatory B. Perhaps A and B are the two (equally) best places to study the clarinet and the two (equally) worst places to study the violin. In which case, Lisa's experience is probably better than Bart's. This example, though, runs foul of the 'other things being equal' clause in claim 3. In my example, Lisa's experience of bliss is just the same as Bart's would have been, had he chosen it. (See the previous note.) In the current example, Lisa's experience of studying the clarinet at B (one of the best places for that purpose) is clearly not the same as Bart's experience of studying the violin at B (one of the worst places for that purpose) would have been.

Life for headaches: there is some finite number of headaches, such that it is permissible to kill an innocent person to avoid them.

A reader of this article, until this point, could be forgiven for assuming that they were reading an argument *against* consequentialism. However, in this section, I will argue that *life for headaches* is not, after all, particularly unpalatable. In fact, I will claim, most of us, consequentialists and nonconsequentialists alike, accept at least some other claims that do not differ significantly from *life for headaches.*

Thousands of people die in automobile accidents every year in the United States. It is highly probable that the number of deaths is positively correlated with the speed limits in force on highways, at least within a certain range. One of the effects of raising speed limits is that there are more accidents, resulting in more deaths and injuries. One of the effects of lowering speed limits is that there are fewer accidents. Higher vehicle safety standards also affect both the numbers of accidents and the severity of the injuries sustained when accidents do occur. Another effect of raising speed limits is that more gasoline is consumed, which raises the level of particulate pollution, which also leads to more deaths.[23] Stricter standards for fuel efficiency also affect the amount of gasoline consumed. There are, then, many different measures that we, as a society, could take to lower the number of automobile-related deaths, only some of which we do take. There are also many measures we could take, that would *raise* the number of such deaths, some of which we do take. Furthermore, it is not obvious that we are wrong to fail to do all we can to reduce the number of deaths. For the purposes of this discussion, I will focus on just one aspect of this failure, the failure to impose a national speed limit of 50 miles per hour in the US.

If there were a national speed limit of 50 mph, it is overwhelmingly likely that many lives would be saved each year, as compared with the current situation. One of the costs of the failure to impose such a speed limit is a significant number of deaths. The benefits of higher speed limits are increased convenience for many. Despite this, it is far from obvious that the failure to impose a 50 mph speed limit is wrong.[24] In fact, most people believe what I will call *lives for convenience:*

23. According to a study by the Natural Resources Defense Council, "Some 64,000 Americans may die prematurely each year because of air pollution" (Reuters: May 9, 1996).

24. There are those who react to what I say about the 50 mph speed limit by declaring that I have convinced them that it *is* wrong not to impose it. But what I say about the 50 mph speed limit can also be said about a 40 mph speed limit, or a 30 mph speed limit,

Lives for convenience: We are not morally obligated to impose a national speed limit of 50 mph (or less).

If we reject *life for headaches* as obviously wrong, we must find a morally significant difference between it and *lives for convenience.* I will devote the remainder of this article to examining the most promising candidates for such a difference.[25]

(i) The victims of higher speed limits are unknown, whereas *life for headaches* envisages the death of an identifiable victim. Even if we could tell how many more people die because of higher speed limits, we could not tell which of the victims would not have died, had the limit been lower. We cannot, therefore, point to specific individuals as the ones whose lives were the price for our increased convenience.

It is hard to see how this epistemic difference could make a moral difference. Consider the following two scenarios: *A:* New York City is struck by a series of freak accidents involving power plants, which result in the deaths of almost everyone within a twenty-mile radius of the Empire State Building. The governor is informed that there are a handful of survivors. He orders the national guard to round up the survivors, make a record of their names, and kill them. *B:* In response to the same information as in the previous example, the governor orders a hydrogen bomb to be detonated in New York City, thereby eliminating the possibility of discovering who survived the initial accidents. I hope it is clear that at least part of what is morally wrong with the governor's actions in each case is that some people are killed who would otherwise have lived. It makes no difference that in case *B* it is impossible to tell which people were killed by the nuclear explosion.

(ii) It is not just that we cannot tell whose deaths are attributable to

---

or even about abolishing private automobiles altogether. Very few are hardy enough to follow their respect for life to such extremes.

25. The discussion that follows is similar, in some respects, to the more detailed discussions in Peter Unger's recent book, *Living High and Letting Die* (Oxford: Oxford University Press, 1996). We are both concerned to identify factors that might account for different intuitive judgments on a pair of cases. Unger's cases involve allowing considerable harm to befall others at the benefit of increased convenience, or even just avoiding minor expense, of one. Unlike my examples, Unger's do not address the possibility of aggregating minor benefits to outweigh major harms. Unger is also more concerned than I am to identify psychological factors that influence differential judgments. One of his aims is to strip such factors of their force by exposing them. I only discuss factors that have at least an outside shot at being morally relevant. Our overall aims are similar, though. We both argue that the initial intuitive judgments about our cases are mistaken.

the higher speed limits, it is that there is no fact of the matter as to who would not have died, had the limits been lower. A counterfactual of the form 'x would not have died had the speed limit been lower' has no determinate truth value. Even if this is not so, it is, at least, indeterminate who *will* die as a result of raising or not lowering the speed limit.

The weaker of these claims clearly has no moral relevance. Consider these two prospective laws, to be considered by the Texas State Legislature: *I:* two victims are to be picked at random from the electoral roll one day after the passage of this law. They are to be executed two days later. *D:* Fred and Barney, whose names were picked at random from the electoral roll a few minutes ago, are to be executed three days after the passage of this law. The fact that, at the time of passage, it is indeterminate who will die as a result of *I*, but determinate who will die as a result of *D*, obviously makes no moral difference.

Even if there is *never* any fact of the matter as to whose deaths are the result of higher speed limits, this makes no moral difference. Consider the following scenario: Two prisoners of conscience, Smith and Jones, are slated for execution in a small totalitarian republic. The president, Shrub, troubled by the effect of an Amnesty International campaign on his public image, decides to spare one of the prisoners. He cannot make up his mind whose life to spare, so he devises the following apparatus: Smith and Jones are placed in separate cells, each with air vents leading to a canister of poison gas, which is set to release its contents at noon. A computer is programmed to select a three-digit number at random at one second before noon. If the number is even or zero, the computer will close the air vent in Smith's cell; if the number is odd, the computer will close the air vent in Jones' cell. The random number selection process is truly indeterministic. The vice president of the republic, Fowl, does not approve of Shrub bowing to liberal pressure. Fowl unplugs the computer at one minute before noon. The gas is released at noon, and both Smith and Jones die. I hope it is clear that Fowl has done something bad in this example. Two people have died instead of one. It is also clear that there is no fact of the matter as to which prisoner has been killed by Fowl's actions. Since the number selection process is truly indeterministic, there is simply no fact of the matter as to whether the number would have been even or odd. There is no nonarbitrarily identifiable victim of Fowl's action, and yet it is morally on a par with killing one person.

Even though neither (i) nor (ii) can ground a *moral* difference between *life for headaches* and *lives for convenience*, they might explain a *psychological* difference between them. We are more inclined to give to charities when we are told of specific individuals who will benefit from our gift. When the story of a child, whose parents cannot afford his life-saving operation, is aired on television, complete with interviews with the child and parents, donations flood in. When we are told of a natural disaster, that threatens the lives of thousands of children, our donations may even be less than in the previous case, and they will almost certainly not be thousands of times greater. Clearly, however, the life of the one child is no more morally significant than the lives of any of the others. If consequentialism's commitment to *life for headaches* is to be a strike against the theory, we must do better than (i) or (ii).

(iii) Speed limits are set by the government, which is elected by the people. So higher speed limits are the end product of a democratic process, in which the victims participate. In a sense, then, the victims of higher speed limits have chosen the system that kills them. This is not true of the prospective victim of *life for headaches*.

First, many of the victims of higher speed limits do not participate in the democratic process. In the 1994 US congressional elections, only 59 percent of registered voters participated, and only two-thirds of eligible voters were registered. Thus, only about 39 percent of eligible voters participated. But, it might be objected, the other 61 percent could have participated, had they chosen. True, but many of the victims of higher speed limits are not even eligible to vote, since they are children. Second, the question of speed limits was not a campaign issue, so many voters who did participate did not know their chosen candidate's stand on the issue. Third, it is also possible that many voters had to choose between candidates, *all of whom* supported higher speed limits, so they did not even have the option to vote against them. Fourth, it is highly probable that some of the victims of higher speed limits *did* vote for candidates who opposed such limits.

(iv) Even though (at least some) victims of higher speed limits cannot be said to have *chosen* the system that kills them, at least they have *benefited* from the system, along with all the other automobile users.

Once again, this is most likely not true of all victims. Some may be drivers or passengers in vehicles that do not take advantage of higher speed limits. Some may be pedestrians. Some may never even approach

an automobile, and yet be killed by the higher levels of particulate pollution caused by the greater consumption of gasoline that results from higher speed limits. Besides, our intuitive reaction to *life for headaches* would not change if it were specified that the victim was one of the many headache sufferers, and that her headache would be relieved before she were killed.

(v) The victims of higher speed limits, even though they may have neither chosen nor benefited from those limits, have at least freely chosen to undertake the risk of being harmed. The dangers of driving, or being driven, are well known. Those who choose to travel by road, therefore, are at least partly responsible for any harm that befalls them. The same cannot be said of the prospective victim of *life for headaches*.

Once again, this is clearly not applicable to the victims of higher levels of particulate pollution. Nor does it seem to be true of *all* the victims of road accidents. Many children are killed on the roads. Many of these may have had no say over whether they were to travel that way. Perhaps we will say that their parents voluntarily assumed the risk on their behalf. But this seems to be an inadequate reply. Would our reaction to *life for headaches* be significantly different, if the prospective victim were a child, chosen at random from among those whose parents had agreed to the selection procedure? Besides, it is not clear how *free* is the choice to travel by road, even for well-informed adults. For many of the victims of road accidents, the alternatives may have been excessively burdensome, if not nonexistent. Many people do not have access to basic services, such as groceries and health-care providers, except by road.[26]

(vi) We do not know for certain that more people die on the roads (or from particulate pollution) when speed limits are higher. The evidence, though inductively strong, is not conclusive. Therefore, there is *some* chance, albeit small, that a decision to raise (or not lower) speed limits will not result in more deaths. Even if the alternatives being considered were the total abolition of private cars versus the status quo, it is *possible* that the former would not save any lives over the latter. *Life for headaches*, however, deals with the death of an innocent, not simply the (overwhelming) likelihood of death.

Certainly it is *possible* that higher speed limits will not result in more

26. For an interesting discussion of this point, and others relevant to the current discussion, see Guido Calabresi, *Ideals, Beliefs, Attitudes, and the Law* (Syracuse: Syracuse University Press, 1985), Chap. 1.

deaths than lower ones. Perhaps drivers drive more carefully at high speed, or car manufacturers work harder to develop safety features when speed limits are higher. It is even possible that the abolition of private cars would not save lives. The chances, however, are exceedingly slim. Can the difference between certainty and near certainty of death really mark a morally significant difference between *life for headaches* and *lives for convenience*? I think not. The seeming counterintuitiveness of *life for headaches* would hardly be diminished if we specified that the victim were merely to be shot in the head at close range with a Magnum .44. There would, in that case, be a small, but finite, probability that she would not die. Or perhaps the victim could be made to play Russian roulette with a thousand-barreled revolver with only one empty chamber. Such alterations do not change most people's intuitive reactions to *life for headaches*.

(vii) When we raise (or fail to lower) speed limits, we do not *kill* anyone. The most we do to the victims of higher speed limits is to *let them die*. *Life for headaches* envisages *killing* an innocent victim. Since there is a morally significant difference between killing someone and letting them die, there is a morally significant difference between *life for headaches* and *lives for convenience*.

Whether the distinction between killing and letting die is morally significant is highly controversial. There is barely any agreement on what the distinction amounts to, let alone whether it can bear moral weight.[27] However, let us assume, for the sake of argument, that we understand what the distinction is and that it is morally significant. Have we now found the difference that allows us to say that *life for headaches* is obviously wrong, whereas *lives for convenience* is not wrong? That depends on just what the moral significance of the killing/letting die distinction is supposed to be. Perhaps it *is* worse to kill someone than to let them die, but only *proportionately* worse. Perhaps, that is, it is as bad to kill one person as to let, say, ten people die, or twenty, or a hundred. On the other hand, perhaps it is *lexically* worse to kill than to let die. That is, it is worse to kill even one person than to let *any number* of people die. Only if killing is lexically worse than letting die might the killing/letting

27. For more on these questions, see Bonnie Steinbock and Alastair Norcross, eds., *Killing and Letting Die* (New York: Fordham University Press, 1994), and Jonathan Bennett, *The Act Itself* (Oxford: Oxford University Press, 1995).

die distinction do the work we need it to do. Suppose it is only proportionately worse to kill than to let die. If we can justify letting someone die by the resulting gains in convenience to many, we can justify killing someone by the resulting gains to proportionately more. If the significant difference between *lives for convenience* and *life for headaches* is that the latter involves killing, but the former involves letting die, and *lives for convenience* is correct, then we *can* justify letting some people die by the resulting gains in convenience to many. So, just how defensible is the claim that killing is lexically worse than letting die? If it is true, it follows that it is worse to kill one person than to let the entire sentient population of the universe (including the one person) die. This is clearly unacceptable.

(viii) When we raise (or fail to lower) speed limits, we do not *intend* any of the deaths that we thereby bring about, even though we may *foresee* them. *Life for headaches* envisages the *intentional* killing of an innocent victim. Since there is a morally significant difference between bringing about a death that is intended and bringing about a death that is merely foreseen,[28] there is a morally significant difference between *life for headaches* and *lives for convenience.*

Once again, the questions of what the intending/foreseeing distinction is and whether it can bear moral weight are controversial.[29] Let us assume, again, that we understand the distinction, and that it is morally significant. It is clear, for the reasons given above with respect to the killing/letting die distinction, that bringing about a death that is intended will have to be lexically worse than bringing about one that is merely foreseen, if the distinction is to mark the right sort of difference between *lives for convenience* and *life for headaches.* But, once again, this is an untenable position. Could it really be worse to bring about and intend the death of one person than to bring about, but merely foresee, the deaths of everyone, including the one?

There is another possibility to consider. Even though killing may not

28. The moral significance of this distinction is often defended in the context of the Doctrine of Double Effect, according to which the distinction between evil that is an intended means to a good end and evil that is a foreseen side effect of achieving a good end is crucial.

29. For more on these questions, see, among others, Steinbock and Norcross, *Killing and Letting Die,* esp. Chaps. 6–9, 20, and Introduction to the 2nd ed., and Bennett, *The Act Itself.*

be lexically worse than letting die, and intending a death may not be lexically worse than foreseeing it, perhaps intentionally killing may be lexically worse than merely foreseeing the death of someone you only let die. Since, it may be argued, we neither intend the deaths of the victims of higher speed limits, nor do we kill them, we must consider the two distinctions working together. This suggestion, however, is just as implausible as the previous ones. Consider the following scenario:

> *Jane, Dick, and the fate of humanity:* Jane is a crack shot with a high-powered rifle, who has just discovered a nefarious plot against all humanity. A bomb, powerful enough to destroy all life on Earth, is connected to an ingenious trigger device. The device, hidden somewhere in Disneyland, is set to detect any signs of human life within the amusement park. If it detects any signs at noon, it will detonate the bomb. It had been thought that the park was clear, but, at the last minute, Dick, who is absent-minded and hard of hearing, is discovered several hundred yards inside the perimeter. It is now ten seconds before noon, and the only way to prevent the bomb exploding is for Jane to shoot and kill Dick with her high-powered rifle. If she shoots him, she will intentionally kill him. If she does not shoot, she will foresee the deaths of all humanity, whom she will let die.

Clearly, it cannot be worse for Jane to shoot Dick than to allow all living things on Earth, including Dick and herself, to be killed by the bomb, even though she would be neither killing them nor intending their deaths.

There is another reason why even a combination of the killing/letting die distinction and the intending/foreseeing distinction cannot ground the right sort of difference between *lives for convenience* and *life for headaches.* A version of *life for headaches*, in which the victim is let die and her death is merely foreseen, would not seem significantly less counterintuitive, at least from the point of view of common-sense morality. Consider this scenario:

> *Drowning baby:* Fred hears of an outbreak of minor headaches, lasting from 2 until 3 every afternoon, in a neighboring town. He happens to possess the only known cure, which, if he gets it to the town before 2, will prevent all the headaches. Otherwise the people will suffer for an hour. Fred is hurrying to catch the only train that will get him there in time, when he passes a baby drowning in a pond. If he stops to

rescue the baby, which he can accomplish easily, he will miss the train, and all the people in the neighboring town will suffer a minor headache for one hour. If he does not stop, the baby will die, but Fred will neither have killed her nor intended her death.

Now consider the following claim:

Life for headaches 2: there is some finite number of headaches, such that it is permissible for Fred to abandon the baby to a watery grave in order to avoid them.

I think it is clear that whatever counterintuitiveness attaches to *life for headaches* also attaches to *life for headaches 2.* At the very least, it is hard to see how one could judge the former to be obviously wrong and the latter to be right.

Despite the seeming counterintuitiveness of *life for headaches* and the inoffensive nature of *lives for convenience,* I have been unable to find a difference between them that would justify the claim that one is false and the other true. There may yet be such a difference, but I think I have looked hard enough, for someone who does not believe *life for headaches* to be false. I now pass the burden of discovering such a difference to those who stubbornly persist in believing that *life for headaches* is not only counterintuitive, but false. In the meantime I conclude that consequentialism's commitment to *life for headaches* is not a decisive strike against the theory.