Counterexamples to the Transitivity of *Better Than*

Stuart Rachels

I. Why the Thesis is Not Too Ridiculous to Take Seriously

Ethicists and economists commonly assume that if A is all things considered better than B, and B is all things considered better than C, then A is all things considered better than C. Call this principle *Transitivity*. Although it has great conceptual, intuitive, and empirical appeal, I will argue against it.

Some philosophers believe that the concept of *better than* implies Transitivity, so Transitivity must be true.[1] But this is no more compelling, without further argument, than saying that the concept of *time* implies absolute simultaneity, and so absolute simultaneity must be true.

Our persistent, forceful belief in Transitivity goes with the idea that value is like a line. If the value of outcomes or states of affairs can be represented as points along a line, then *better than* seems transitive because *to the right of* seems transitive. However, the linear view of value is not sacrosanct. Twentieth century physics has questioned Kant's analogous idea that time is like a line.

Our only empirical evidence for Transitivity is inductive: we have frequently observed that some A is better than some B which is better than some C which is worse than A. However, even if all known instances conform to Transitivity, some unusual cases may violate it. Larry S. Temkin explains how three types of ethical principle, which cannot be dismissed *a priori*, threaten Transitivity: (a) principles implying that in some cases different factors are relevant to comparing A to C than to comparing A to B or B to C; (b) principles of limited scope; (c) principles implying that morally relevant differences in degree can

amount to differences in kind. I won't rehearse Temkin's explanations for why principles of types (a) and (b) threaten Transitivity, but they do.[2] My counterexamples to Transitivity employ a principle of type (c): pleasures and pains enormously different in intensity differ in kind. Temkin endorses this type of counterexample, using arguments based on earlier drafts of this paper.[3]

II. Four Preliminary Points

I should make the following points explicit before giving the counterexamples:

1. Transitivity is formulated in terms of the relation *better than*. It stands or falls with the corresponding principle for *worse than*.
2. Although Transitivity refers to only three outcomes—A, B, and C—cases involving more outcomes can contradict it. It can easily be proved that a set of outcomes violates Transitivity if those outcomes can be ordered so that each subsequent outcome is worse than the one prior even though the last outcome is not worse than the first.
3. The outcomes in my examples have implicit *ceteris paribus* clauses. I will say that one outcome is better than another based on incomplete descriptions; these descriptions may be amplified only such that the added details do not alter the overall value of the outcomes.
4. These examples employ stipulated, technical concepts of *pleasure* and *pain*: F's total conscious state is pleasurable just in case that experience, considered merely as feeling, is preferable to temporary unconsciousness; F's total conscious state is painful just in case temporary unconsciousness is preferable to that experience, considered merely as feeling. Stipulating that *pleasure* and *pain* be understood in these ways commits us to nothing substantive and facilitates the articulation of the examples. The most important difference

between these concepts and the ordinary ones is that mild pains, on this view, are more unpleasant than the sensations we normally call mildly painful. Ordinarily we say that a hangnail is mildly painful, but having a hangnail is not unpleasant enough for us to call it "painful" here because it is usually not worse than temporary unconsciousness. Also, what we normally think of as "mild pleasures" feel better than sensations barely better than temporary unconsciousness. In fact, many of the sensations we will here call "mildly pleasurable" would ordinarily be called "painful," since we prefer some unpleasantness to anesthesia. If we imagine a left-to-right continuum of sensation, with ecstasy on the far left and agony on the far right, then the main impact of our stipulated conceptions is that the range of pleasures expands to the right, reducing the range of pains. We should keep an eye on these conceptions, so my arguments never gain mileage by conflation with ordinary usage.

III. The First Counterexample: Comparing Long Periods of Pain

Mild pains are barely worse than temporary unconsciousness. To focus our intuitions, here are two experiences that many people find mildly painful:

—Hospital patients being prepared for surgery are sometimes given the option of being sedated instead of being wheeled into the operating room and anesthetized there. Many patients accept this offer, presumably because they believe that temporary unconsciousness is better than the anxious experience of being taken into surgery and put on the operating table.
—Many people who anticipate working the entire evening on a tedious project due on the following day would prefer to snap their fingers and wake up the next morning, the project completed, rather than spend those hours working. Many people have such a preference

because they consider temporary unconsciousness better than spending several hours laboring over an uninspiring task.

The first counterinstance to Transitivity consists in the outcomes A-Z. Z consists in a tremendously long period of time, each moment of which is slightly worse than temporary unconsciousness. For simplicity one might think of Z as consisting in many years of being wheeled into surgery or writing school reports researched in the *World Book Encyclopedia*. In thinking about Z do not imagine that each time you write a report, you anticipate writing millions more and can only remember writing millions before. For if that were the case, your life would be hellish, and each moment would be far worse than temporary unconsciousness. So, again to simplify, it may be helpful to imagine Z as a series of mildly painful experiences in which you neither anticipate nor remember similar episodes.

The first counterexample depends on the idea that A is worse than Z. A is horrible agony for one year. The pains in Z and A *differ in kind* in the sense that no finite duration of Z is as bad as a year (or so) of A. To make this rhetorically compelling, I should now describe a method of excruciating torture that would hint at the enormity of A. (Wouldn't you rather be wheeled into the operating room repeatedly than have Satan do X, Y and Z to you for a year?) However, since torture is such an unpleasant subject, I leave this task to your imagination, if you doubt that Z is better than A.

Each outcome in the example involves a single person's experience:

A: 1 year of excruciating agony.

B: 100 years of pain slightly (or somewhat) less intense than the pain in A.

C: 10,000 years of pain slightly less intense than the pain in B.

D: 1 million years of pain slightly less intense than the pain in C.

. . .

Y: 10^{48} years of pain slightly less intense than the pain in X.

Z: 10^{50} years of pain slightly less intense than the mild pain in Y.[4]

What happens to the poor suffering souls after the year in A, the 100 years in B, and so on? I do not think it matters, if everyone in A-Z has the same fate. Two variations are *death* and *normal life*.

Although A is worse than Z, the example creates a path from A to Z involving only changes for the worse. These changes are for the worse because increasing a pain's duration 100-fold offsets reducing its intensity slightly, or even somewhat. So the outcomes get worse until they are better, contradicting Transitivity. B is worse than A, C is worse than B, D is worse than C . . . and Z is worse than Y, yet Z is better than A.

Now I will respond to the four most plausible objections to this counterexample.

The first holds that this is a Sorites Paradox. Sorites Paradoxes are unsound arguments. However, this counterexample is not of the Sorites type. Sorites arguments appeal to a series of steps each of which makes no difference to applying a concept. For example, having one hair fewer makes no difference to whether someone is bald. In my example, each step and each change makes a difference. Each outcome is worse than the one before it; each change in the intensity of a pain makes that pain slightly more intense, and thus worse.[5]

The second objection is that the fanciful nature of the example undermines its credibility. This reply might rely on either of two theses. The first is that our judgments about incredibly long, painful lives cannot be trusted. The second is that theoretical concepts (such as *better than*) only answer to physically possible outcomes, and many of the lives in this example are impossible given the laws of nature.

These objections, even if sound, would not discredit the third counterexample (below), which involves much shorter durations. However, it is not successful even here. In response to the first thesis, most of the people with whom I have spoken have marked opinions about which outcome is preferable in each instance. Although the outcomes are bizarre, each comparison involves only two considerations: intensity of pain and length of time. Thus, we may have informed beliefs in each case. As for the second claim, one may live in pain for millions of years if vastly improbable quantum events keep one's nervous system (and other parts of the world) in proper states. Theoretical concepts such as *better than* should not exclude such possibilities. This is consistent with saying that, for everyday practical purposes, we might reasonably assume that our comparative assessments will accord with Transitivity.

The third objection holds that A is better than Z. According to this response, we have difficulty conceiving how much badness can accumulate, bit by bit, over $1x10^{50}$ years; a sufficient duration of mild pain is worse than a year of wrenching agony. This claim can be defended by a clever argument:

P1) One million years of mild pain are worse than three seconds of agony.
P2) One year of agony (=A) is just 10 billion three-second periods of agony strung together.
C1) Therefore, 10 billion million years of mild pain (10^{16}) are worse than one year of agony (=A).
C2) So Z is worse than A, since 10^{50} years of mild pain (=Z) is even longer than the 10^{16} years of mild pain which is worse than A.

To discuss this argument in detail would introduce issues excellent at dividing opinion. But perhaps we can agree on one weakness. P1 holds that no agony is so

wrenching that three seconds of it are worse than a sufficiently long duration of mild pain. Is there no agony so extreme? Bear in mind, again, that the mild pain does not become worse over time because the person loses hope and feels demoralized; that tendency must be stipulated away or counterbalanced. Edmund Gurney says there is such agony; Henry Sidgwick says there is not.[6] The best way to investigate whether humans have experienced such pain is to interview thoughtful people who have recently suffered. Many people, I predict, would rather feel mild pain for untold aeons than their worst pains for three seconds. And many of those people will stick with this preference even if asked not to discount mild pains because they are in the farther future. If so, then we should reject the third objection. However, even if reliable parties prefer the three seconds of agony, the objection fails if physically possible pains more intense than anyone has experienced are worse than any duration of mild pain. That possibility is highly likely.

The fourth and most tempting objection is that at least one of our other judgments is incorrect. According to this response, it is better to experience some slightly less intense pain rather than a slightly more intense pain, even though the less intense pain lasts 100 times longer. This claim is tempting only for pains that are mild: one might hold that it is good for a mild pain to become a more mild pain with a duration 100 times greater. However, no one believes that it is all good, on the whole, for a severe pain to become slightly less intense but 100 times as long.

The fourth objection ignores our stipulated conception of pain. The judgment on which it rests is plausible only if we mistakenly conceive “mild pain” as being better than, or about as good as, temporary unconsciousness. But if we bear in mind that even the mildest pains are at every moment worse than temporary unconsciousness, then we should prefer to increase their intensity slightly if that would reduce our pain sentence by 99%. For example, we would prefer being slightly more anxious on the way to the operating room rather than

repeating the slightly less anxious experience 100 times. If the objection still seems forceful, suppose it claims that V is not better than W. V is fairly mild pain for many years; W is slightly less intense pain for 100 times as long. If so, the first example still violates Transitivity if V is better than A: for then B is worse than A, C is worse than B, D is worse than C, etc., yet V is better than A; the outcomes get worse until they are better. And V is better than one year of agony because the pains in V are fairly mild.

Other objections to the first counterexample seem obviously inadequate, so I will discuss them in an endnote.[7]

IV. The Second Counterexample: Comparing Long Periods of Pleasure

The second counterexample to Transitivity is analogous to the first. The difference is that the outcomes involve pleasure rather than pain. Z is many years of the mildest pleasures, or conscious states barely preferable to temporary unconsciousness. We may think of Z as consisting in less unpleasant versions of our mild pains, that is, being wheeled through the hospital in somewhat better spirits or working on a task that is only somewhat burdensome. Or, to use Parfit's memorable example, we may think of Z as a life of the barest pleasures: muzak and potatoes.[8] Z is worse than A, which is 50 years of the most intoxicating joy.

A: 50 years of ecstasy.

B: 5,000 years of pleasure slightly less intense than the pleasure in A.

C: 500,000 years of pleasure slightly less intense than the pleasure in B.

D: 50 million years of pleasure slightly less intense than the pleasure in C.

. . .

Y: $5x10^{49}$ years of pleasure slightly less intense than the pleasure in X.

Z: $5x10^{51}$ years of pleasure slightly less intense than the mild pleasure in Y.

In these outcomes B is better than A, C is better than B, D is better than C . . . and Z is better than Y, but Z is worse than A. The outcomes get better until they are worse, so Transitivity is violated. Objections analogous to those above apply here. We need not reconsider objections one and two (the Sorites objection and the Fanciful Nature objection) because our assessments of them would not change.

Objections three and four should be reconsidered. They concern the comparative value of outcomes. Given that pain and pleasure may be disanalogous in various respects, these objections might prove more powerful against the second counterexample than against the first.

According to the third objection, our inability to conceive vast durations distorts our judgment: Z is better than A. We do not realize how much goodness can accumulate, bit by bit, over $5x10^{51}$ years; a sufficient duration of mild pleasure is better than fifty years of ecstasy. But I have found few people who would prefer aeons of the drabbest pleasures to a lifetime of bliss. Derek Parfit, for example, says that he would prefer "the Century of Ecstasy" to "the Drab Eternity."[9] The Drab Eternity offers only the pleasures of muzak and potatoes. Though meager, muzak and potatoes can be imagined to offer a life that is at each moment better than temporary unconsciousness.

If you accepted the third objection to the *first* counterexample, then you may be less tempted here. That objection claims that a sufficient duration of mild pain is worse than a year of wrenching agony. If that claim persuades you, then you are probably impressed by the awful nature of mild pain. And if you are so impressed, then you may be correspondingly unimpressed by the tepid nature of mild pleasure, since mild pleasures and

mild pains are near neighbors on the hedonist's continuum. For that reason you may agree with me that fifty years of ecstasy are better than a lengthy duration of meager pleasure.

The third objection, however, can be defended as follows:

P1) One million years of mild pleasure are better than three seconds of bliss.

P2) 50 years of ecstasy (=A) are just 500 billion three-second periods of bliss strung together.

C1) Therefore, 500 billion million years of mild pleasure ($=5x10^{17}$) are better than fifty years of ecstasy (=A).

C2) So Z is better than A, since $5x\ 10^{51}$ years of mild pleasure (=Z) is even more than $5x10^{17}$ years of mild pleasure which is better than A.

This argument raises controversial issues.[10] But perhaps we can agree (in analogy to the first counterexample) that three seconds of the most intense ecstasy are better than any duration of the mildest pleasures. Dostoevsky wrote about such sensation in a letter:

> In certain moments, I experience a joy that is unthinkable under ordinary circumstances, and of which most people have no comprehension. Then I feel that I am in complete harmony with myself and the whole world, and this feeling is so bright and strong that you could give up ten years for a few seconds of that ecstasy—yes, even your whole life.[11]

The third objection fails because a sufficient difference in degree makes pleasures different in kind.

Objection four claims that it is unwise to trade one pleasure for another that is slightly less intense but 100 times as long. Again, such an objection is only plausible for the mildest pleasures, and even then only if we mistakenly suppose that some pleasures are at least as

bad as temporary unconsciousness. For although we should not trade experiences of small value for longer experiences of no value, we should trade experiences of small value for experiences 100 times as long with less, but at each moment some, value.

V. The Third Counterexample: Twelve Bad Headaches

The third counterexample involves headaches of various intensities and durations:

A: 5 minutes: a wrenching migraine headache. Your head is ready to explode.

B: 10 minutes: a pounding migraine headache somewhat less bad than the headache in A.

C: 20 minutes: a hideous headache somewhat less bad than the headache in B.

D: 40 minutes: a terrible headache somewhat less bad than the headache in C.

E: 90 minutes: a dreadful headache somewhat less bad than the headache in D.

F: 3 hours: a headache somewhat less bad than the headache in E.

G: 6 hours: a headache somewhat less bad than the headache in F.

H: 12 hours: a headache somewhat less bad than the headache in G.

I: 1 day: a headache somewhat less bad than the headache in H.

J: 2 days: a headache somewhat less bad than the headache in I.

K: 4 days: a headache somewhat less bad than the headache in J.

L: 1 week: a headache somewhat less bad than the headache in K. Its pains are only slightly worse than temporary unconsciousness.

As we move down the alphabet, the headaches get all things considered worse because having a painful headache is worse than having a headache somewhat more painful for only half as long. Transitivity is violated because A is worse than L yet it becomes L via only changes for the worse.

To accept this counterexample we must answer Yes to two questions. First, can five minutes of a severe, agonizing migraine headache be worse than a week of mild pain?

One reason to answer No is that it is worse to lose a full week of normal life rather than five minutes. Our example, however, should not be susceptible to this objection because we want to compare the pains alone. We should stipulate that if one has a headache for less than a week, then one will spend the rest of the week unconscious, or at least having a life neither better nor worse than unconsciousness and no more productive than if one had the one week headache.

People who have had severe migraine headaches are most qualified to judge whether a very bad migraine is worse than a week of mild pain. In conversation I have found that some such people believe that five minutes of a bad migraine headache are worse than a week of mild pain. Thus, we should answer Yes to the first question.

Second, can a severe headache be transformed into a mildly painful headache with 11 moderate reductions of intensity? Making a pain somewhat less intense eleven times should make its intensity very much less severe and thereby bridge the gap between the extremes. Eleven steps may be more than we need. Again we should answer Yes.

If we accept these two answers, then we must deny Transitivity.

Before we move on, notice that we could easily construct an example analogous to this one involving pleasure. Outcome A, five minutes of the most heavenly sexual pleasures, is better than outcome L, a week of lousy sex. (If Nagel is right that bad sex is better than no sex, then lousy sex should be better than temporary unconsciousness.[12]) L could be reached from A via 11 moderate reductions in pleasure intensity, thus ensuring that B would be better than A, C better than B, and so on. In general, one can formulate variations of the first three counterexamples by altering a) the types of pain or pleasure involved, b) their durations, and c) the degree to which intensities change across outcomes.

VI. The Fourth Counterexample: Parfit's Second Paradox

Creative work in philosophy typically arises out of existing ideas; so what one person does would most likely have been done by someone else, given a few more years. But Derek Parfit has devised a problem so clever and original that it may be an exception to this rule. Parfit calls this "the Second Paradox."[13]

The Second Paradox is a set of outcomes ordered so that the outcomes seem to get better and better, yet the last outcome is worse than the first. The paradoxical conclusion, derived with Transitivity, is that the last outcome is better than the first. Parfit does not reject Transitivity; he resolves the paradox differently. I will argue that Parfit's and Temkin's resolutions of the paradox are inadequate, and that the Second Paradox is a counterinstance to Transitivity.

The Outcomes in the Second Paradox

The Second Paradox is tedious to work through. However, once grasped, it dazzles the philosophical imagination.

The first outcome is A+. (See the diagram.) A+ contains two groups of 10 billion people: one group whose lives are at 100, an ecstatic level, and another whose lives are at 50, a level of pleasure well worth enjoying. The last outcome is Omega 100, a world that contains many, many lives each of which is barely worth living at each moment. In Omega 100 muzak and potatoes are the only pleasures in life. Although A+ is better than Omega 100, A+ is transformed into Omega 100 via only changes for the better.

Each change from A+ to Omega 100 takes one of two forms.

The first kind of change occurs as A+ becomes Alpha. This happens by raising both groups in A+ to a 105 level of pleasure and adding many, many groups of 10 billion people whose lives, at 45, are well worth living. The outcome is much improved going from A+ to Alpha because all the people in A+ benefit from the change, especially those in the 50 group, and the only 'cost' of this benefit is adding people to the world who are glad to be alive.

The second kind of change occurs as Alpha is transformed into Beta. This occurs by lowering the two better-off groups in Alpha from 105 to 104 but raising as many worse-off groups from 45 to 104. (Even after this change there are many groups at 45.) This kind of change occurs down the Greek alphabet until we reach Omega. In Omega, there are many groups at 90 but many still at 45.

Omega is transformed into Alpha 2 by improving all the lives in Omega to 95 (including the lives that were at 45) and adding many more groups at 40. This repeats the first sort of change. Alpha 2 is transformed into Beta 2 by lowering the better-off groups to 94 but raising the same number of worse-off groups to 94. By the time we reach Omega 2,

the better-off groups are down to 80, though there are many more of them, while there are still many groups at 40. At Alpha 3 all the people in Omega 2 are promoted to the level of 85 and many groups at 35 are added.

So at each Omega the average quality of life is lower than it was at the previous Omega, and the population has been greatly increased. At Omega 100, everyone's life is barely worth living at each moment. We want to say both that Omega 100 is worse than A+ and that each change from A+ to Omega 100 is for the better. Each change seems to be for the better because the quality of life is only lowered for those who are better off, and then only when this loss is offset by greater gains for the worse off.

The Second Paradox may also be formulated in terms of painful lives rather than pleasurable ones. To do this, change the numbers in the Second Paradox to negative numbers; then the outcomes will get worse until they are better (the last outcome better than the first), thus contradicting Transitivity.

Parfit's Suggestion

To resolve the paradox, Parfit suggests that Alpha is better than Beta, that is, it is better to have 20 billion people at 105 plus many more people at 45 rather than 40 billion people at 104 plus many more people at 45 (20 billion fewer at 45 than in Alpha).

Parfit defends this by appealing to "Perfectionism." Perfectionism is the view that "even if some change brings a great net benefit to those who are affected, it is a change for the worse if it involves the loss of one of the best things in life."[14] So Alpha is better than Beta because in Alpha the best things are better. In Alpha the luckiest 20 billion listen to Mozart; in Beta, 40 billion listen to Haydn.[15]

However, we assume that the only factor relevant to the quality of life in the Second Paradox is the intensity of the pleasures. We do not assume that the "best things in life"—Mozart's music, for example—are lost as the quality of life drops; the quality of life could fall because the best things are appreciated less. Thus, Perfectionism is only relevant as a defense of Parfit's resolution if we interpret "the best things in life" as the most intense pleasures. Yet on this interpretation, Perfectionism is wildly implausible. It would imply, about our own lives, that three minutes of the best pleasure are better than fifty years of pleasure only slightly less intense. As Parfit admits, "[Perfectionism] conflicts with the preferences most of us would have about our own futures."[16] Furthermore, Perfectionism applied to populations implies that it could be bad for the middle class to become rich at a tiny cost to the wealthy. Hence, we should reject Perfectionism.

Temkin's suggestion

Temkin suggests that many of the moves in the Second Paradox can be blocked by egalitarian considerations. Consider the comparison between A+ and Alpha. The fact that Alpha includes many more worse off people than A+ may represent an egalitarian respect in which Alpha is worse than A+. Perhaps this respect outweighs other factors and yields the result that A+ is all things considered better than Alpha. Or consider the comparison between Alpha and Beta. The inequality between the better off groups and the worse off groups may be worse in Beta because Beta has even more better off people that the worse off can resent (or that we can resent on their behalf). And again, perhaps this consideration suffices to license the judgment that Beta is all things considered worse than Alpha.[17]

Temkin's proposals depend on the idea that inequality is bad apart from other values. Temkin never argues this, but in *Inequality* he offers the "fundamental intuition underlying

egalitarianism—that it is bad, unfair or unjust, for some to be worse off than others through no fault of their own . . .”[18] Do such intuitions apply to the Second Paradox? Is it unfair, for example, that in Beta so many people prosper at the 104 level, while others only prosper at the 45 level? Nothing deters us from supposing that, in our version of the Second Paradox, the different groups of 10 billion people live in different galaxies and cannot know of each other’s existence. In such a scenario, intuitions about fairness and justice do not come into play. Saying that a universe is made worse by the mere fact that isolated populations within it have varying degrees of welfare is to make an ethical judgment too much on the model of certain aesthetic judgments. It would lessen the value of a Picasso (= prospering population) to add a section onto it painted by a lesser artist (= less prosperous population). But it would no more lessen the world’s value that a remote population prospers more than we earthlings than that a non-actual population prospers more than we do. Egalitarian considerations should not affect how we evaluate the Second Paradox. At the very least, they do not outweigh benefits enjoyed by billions of people. There is a treatment of the paradox more palatable than this.

A Different Proposal

We should resolve the Second Paradox by reassessing our concepts rather than changing our ethical judgments. If we deny Transitivity, then we may hold that the outcomes in the Second Paradox become better and better *and* Omega 100 is worse than A+. The Second Paradox, like the first three counterexamples, involves pleasures differing in kind and contradicts Transitivity.

VII. Rationality

If my examples succeed, then it is sometimes rational to prefer (in isolation) X to Y, Y to Z, and Z to X. Robert Nozick explains the "money pump" objection to the idea that such intransitive preferences are rational:

> The idea is that with nontransitive preferences, for example preferring x to y, y to z, and z to x, a person who starts with z can be led to pay a small amount to improve his situation to the y which he prefers to it, another small amount to improve his situation to the x which he prefers to y, and then another small amount to improve his situation to z—the very z he started with—which he prefers to x, thus ending up a net loser.

But, as Nozick says,

> The argument assumes that a person is always willing to act on each individual preference, considered in isolation, and willing to act on each one repeatedly, no matter what he may know about how all of them hang together, no matter how he may foresee his sequential action on individual preferences leading him into just this sort of trouble. This is certainly an implausible assumption.[19]

It is implausible; we should reject it. A rational person with intransitive preferences and adequate information cannot be money-pumped because she rejects the principle, *it is always wise to give up something to get something better*. By denying this principle she avoids both contradiction and poverty. For example, having Z, she may consistently (and wisely) refuse to embrace Y, even though she knows that Y is better than Z.

Objection: she could also embrace Y, but then decline X; or accept X and then stay put. How can she (possibly) determine what to do?

A choice is most rational if the reasons for choosing it are stronger than the reasons for choosing any alternative. For cases that violate Transitivity, we lack a powerful reason for any particular choice: we cannot say *this choice yields an outcome that is best*. If we cannot find an acceptable reason to prefer any choice to any other, then we may conclude that all available *choices* are equally wise. (This does not imply that all available *outcomes* are equally good.) If so, then a rational person could not be pumped of a cent: she would keep what she originally has (whether X, Y or Z), not parting with a penny for something better.

However, there may be sound principles that justify preferring some choices to others in cases that violate Transitivity. These principles would be sensitive to the nature of the outcomes involved. We will not discuss what those principles might be. Suffice it to say, if such principles imply that one outcome among several is most rationally targeted, then a rational person could only be pumped of money until she lands on that choice.

So, I see no reason to accept Parfit's view that the failure of Transitivity implies skepticism about practical reasoning.[20]

One final remark about moral reasoning. Abandoning Transitivity has few practical implications. Because there are few exceptions to the principle, it may persist in our reasoning as a rule of thumb. To avoid using rules of thumb, we must make fewer inferential judgments. Instead of inferring that A is better than C (given that A is better than B and B is better than C), we must compare A and C directly.

VIII. Value

The failure of Transitivity implies that we should not conceive value quantitatively or linearly, except perhaps for some practical purposes. In this respect ethics should not strive to be scientific. Ethics is not, ultimately, a quantitative study.

Even those persuaded that Transitivity is false may have trouble giving up these modes of thought. The quantitative conception of value pervades ethical theory: consider expressions such as "the principle's weight" and "average utility." Even the expressions "more value" and "less value" connote quantities of value that are measurable and may be assigned positions on a number line. But we should not understand "A is better than B" in terms of A having more goodness than B.

NOTES

[1] See, for example, John Broome, *Weighing Goods* (Oxford: Basil Blackwell, 1991), pp. 10-12.

[2] Larry S. Temkin, "Weighted Goods: Some Questions and Comments," *Philosophy and Public Affairs* 23 (1994), pp.361-363, and "A Continuum Argument for Intransitivity," *Philosophy and Public Affairs* 25 (1996), pp. 193-194.

[3] See Temkin, "Weighted Goods: Some Questions and Comments," p. 363, and especially "A Continuum Argument for Intransitivity," sect. 4.

[4] Anyone who doubts that 25 reductions in intensity could turn A's pains into Z's pains (such that B is worse than A, C is worse than B, and so on) may replace "A-Z" with "1-50" or "1-200." However, since the temporal difference between adjacent outcomes is 100-fold, the pains in adjacent outcomes should differ sufficiently in intensity that 26 outcomes are enough.

[5] Temkin ("A Continuum Argument for Intransitivity," sect. 5) rebuts this objection at greater length in a different way.

[6] Edmund Gurney, *Tertium Quid*, vol. I (London: Kegan Paul, Trench & Co., 1887) p. 181; Henry Sidgwick, *The Methods of Ethics*, seventh edition (Indianapolis: Hackett Publishing Company, 1906; repr. 1981) pp. 123-4, fn. 1.

[7] (1) According to the first response, A *is* better than Z, yet one ought to choose Z over A when they are the only options. A is better than Z, it is said, because A compares more favorably than Z to some possible though unavailable alternatives. This strategy could be deployed not only against the judgment that A is worse than Z but against any comparative judgment. However, which of the endless unavailable options are relevant to comparing A

and Z, if a direct comparison of them is insufficient to assess their relative value? And why should we believe that the relevant unavailable possibilities would reverse our initial judgment that Z is better? These questions appear unanswerable. See Temkin, "Intransitivity and the Mere Addition Paradox," *Philosophy and Public Affairs* 16 (1987) pp. 180-183 and his discussion of the same material in "Rethinking the Good, Moral Ideals and the Nature of Practical Reasoning," *Reading Parfit*, ed. Jonathan Dancy (Blackwell Publishers: Oxford, 1997), sect. L.

(2) Temkin entertains the view that we should individuate outcomes so finely that the identity of an outcome changes when 'it' is compared to different alternatives. Thus, my example would not contradict Transitivity because it would be wrong to say *Y is better than Z and Z is better than A*; we could only say *Y is better than Z and Z' is better than A*. However, such a view implies that Transitivity, if the principle is even coherent, applies to nothing: for it would never be the case that A is better than *B* and *B* is better than C. Furthermore, even if one argues in an *ad hoc* way that outcomes should be finely individuated only when they threaten Transitivity, such a proposal saves Transitivity only by admitting that the principle cannot always help us choose among several alternatives when our comparative preferences are clear. This all but admits that Transitivity sometimes fails. See Temkin, "Rethinking the Good, Moral Ideals and the Nature of Practical Reasoning," sect. K.

(3) Temkin mentions another proposal for preserving Transitivity in passing: "More cautiously, one may decide that the concept of "better than" is limited in scope, and that for many cases [i.e., apparent counterexamples to Transitivity] one needs another concept for comparing alternatives that is similar in meaning, but intransitive." (Temkin, "Weighted Goods: Some Questions and Comments," p. 361, fn. 22) This suggestion is dubious both

because it is *ad hoc* and because it needlessly multiplies concepts: a concept that is not transitive can be employed in all our comparative, normative judgments.

(4) James Griffin equates utility with informed desire. So, A is better than B just in case the informed desire for A is greater than that for B. A's being better than B, on this view, is nearly the same as A's being a better choice than B. If so, then Transitivity, it seems, would hold: an informed person would never desire A more than B, B more than C and C more than A, for such a person could be persuaded to give up some money to trade A for C, then be persuaded to give up some money to trade C for B, then be persuaded to give up some money to trade B for A. "Better choice than" is transitive, for its point is to guide action, and we wouldn't know what to do if we were offered A, B and C and thought that A is a better choice than B, B is a better choice than C, and C is a better choice than A.

Griffin's theory of utility, if correct, would imply that Transitivity cannot be violated when all options (A, B and C) are present. However, Transitivity would still fail, for in examples like mine, an informed desire for A would be greater than an informed desire for B *isolated from other alternatives*; and similarly, B would be better than C and C would be better than A.

[8] Derek Parfit, "Overpopulation and the Quality of Life," *Applied Ethics*, ed. Peter Singer (Oxford: Oxford University Press, 1986) pp. 160-161.

[9] *Ibid*., pp. 160-161. Also see James Griffin, *Well-Being: Its Meaning, Measurement, and Moral Importance* (Oxford: Oxford University Press, 1986), p. 86.

[10] Noah Lemos would reject the inference to C1. He affirms the *principle of rank*, according to which "a whole can be a higher good even if none of its parts are themselves higher goods." (*Intrinsic Value: Concept and Warrant* (Cambridge: Cambridge University Press, 1994) p. 57) So, fifty years of ecstasy may be a "higher good" compared with mild pleasure, even though any three second period of that ecstasy is *not*, in comparison, a higher good.

[11] Quoted in Geir Kjetsaa, *Fyodor Dostoevsky: A Writer's Life*, tr. Siri Hustvedt and David McDuff (New York: Viking, 1987) p. 149.

[12] Thomas Nagel, *Mortal Questions* (Cambridge: Cambridge University Press, 1979) p. 52.

[13] Parfit discusses the Second Paradox in *Reasons and Persons* pp. 433-437 and "Overpopulation and the Quality of Life," pp. 156-164. My account of the Second Paradox is taken from the latter essay in which Parfit introduces his proposal for resolving it.

[14] Parfit, "Overpopulation and the Quality of Life," p. 163.

[15] *Ibid*., p. 164.

[16] *Ibid*., p. 164.

[17] Temkin suggested this in correspondence, but his published work also bears on these issues, especially *Inequality* (Oxford: Oxford University Press, 1993), chs. 7 and 9.

[18] Temkin, *Inequality*, p. 290.

[19] Robert Nozick, *The Nature of Rationality* (Princeton: Princeton University Press, 1993) p. 140, fn.

[20] As reported by Temkin, "A Continuum Argument for Intransitivity," p. 209.