

Chapter 5: Counterexamples to the Transitivity of *Being Better Than*

Why the Thesis is Not Too Ridiculous to Take Seriously

Ethicists and economists commonly assume that if A is intrinsically better than B, and B is intrinsically better than C, then A is intrinsically better than C. Call this principle *Transitivity*.ⁱ Transitivity provably stands or falls with the corresponding principle for *intrinsically worse than*, so I will treat them together.

I will offer counterexamples to the transitivity of *being intrinsically better than*. These examples employ more than three states of affairs, but a set of possibilities of any size provably violates Transitivity if it can be ordered so that each succeeding possibility is worse than its predecessor but the last one isn't worse than the first. My examples, if successful, also show the nontransitivity of "being all things considered better than," "being hedonically better than" and "being better for a person than."

In general terms, Transitivity might fail because factors determining how A&C compare differ (or differ in significance) from those determining how A&B and B&C compare.ⁱⁱ In my examples, differences in duration and hedonic intensity are the only relevant factors. Intensity differences always matter, but their significance increases dramatically once they become sufficiently great. Temkin endorses this type of counterexample, using arguments based on earlier drafts of this chapter.ⁱⁱⁱ

Most philosophers strongly believe in Transitivity. This belief may derive force from the idea that value is like a line. If all states of affairs can be represented as points along a line, with better possibilities represented to the right of worse ones, then *being intrinsically better than* seems transitive because *being to the right of* seems transitive. However, this linear view of value is not sacrosanct. Many philosophers already reject it, for example,

those who believe that some items can't be compared in terms of value. Einstein, moreover, discredited an analogous picture of time. On that picture, all events are represented as points along a line, with later events represented to the right of earlier ones. This view entails absolute simultaneity, which Einstein rejected, for two events either are or aren't represented at the same point on the line.

Let me press this analogy further. Conceptual arguments for Transitivity, I think, are no more effective than conceptual arguments for absolute simultaneity. Perhaps the linear view of value has become part of the meaning of value-terms; this might account for Transitivity's conceptual appeal. If so, we should revise those concepts. Similarly, the idea of a universal time might be part of what temporal notions mean (at least for ordinary speakers); this would account for relative simultaneity and the twin paradox seeming somehow incoherent. But physicists reject those definitions.^{iv}

Although Transitivity isn't apodictically certain, inductive evidence supports it: philosophers have often noted that an A is better than a B both of which are better than a C. Any successful argument against Transitivity must outweigh this evidence.

The First Counterexample: Nine Bad Headaches

The first counterexample consists in nine possibilities:

- J: 5 minutes: an agonizing, excruciating migraine headache.
- K: 10 minutes: a pounding migraine headache somewhat less bad than the headache in J.
- L: 20 minutes: a hideous headache somewhat less bad than the headache in K.
- M: 40 minutes: a terrible headache somewhat less bad than the headache in L.
- N: 90 minutes: a dreadful headache somewhat less bad than the headache in M.
- O: 3 hours: a headache somewhat less bad than the headache in N.

P: 6 hours: a headache somewhat less bad than the headache in O.

Q: 12 hours: a headache somewhat less bad than the headache in P.

R: 1 day: a headache somewhat less bad than the headache in Q.

Its pains are slightly worse than temporary unconsciousness.

As we move down the alphabet, the possibilities get intrinsically worse because having a painful headache is worse than having a headache somewhat more painful but for only half as long. The pains in R are still bad: each moment is worse than nothing. Were those pains less bad, then R might not be worse than Q: R's pains would be less intense than Q's, and since the conscious states in R would not be worse than neutral, its extra duration would not count against it. Nonetheless, R is not worse than J. Transitivity is violated because the path from J to R consists in nothing but steps for the worse.

One can construct a similar counterexample involving pleasure. J, five minutes of the best sexual pleasures, is not worse than R, one day of lousy sex, each moment being barely better than unconsciousness. But R could be reached from J with 8 merely moderate reductions in pleasure intensity, thus ensuring that J is worse than K (with its double duration), K is worse than L (with its double duration), and so on.

Can migraine pain be so agonizing that five minutes of it would be at least as bad as a day of R's milder pain? Can an agonizing headache be transformed into a R-level headache with 8 reductions of intensity? Answering "yes" to both questions entails the denial of Transitivity. Let's consider them in turn.

First, can J be at least as bad as R? Think about that question in terms of this one: would you prefer to have the day-long headache or five minutes of the worst headache pain?

You might prefer the shorter headache so you could return to normal life after five minutes rather than a day. Let's stipulate, however, that if one has a headache for less than a day, then one will spend the rest of the day unconscious, or at least having a day neither better nor worse than if one were.

What Parfit calls the *bias towards the future* suggests another motive for preferring the shorter headache. “Looking forward to a pleasure,” Parfit says, “is, in general, more pleasant than looking back upon it. And in the case of pains the difference is even greater.”^v For this reason, Parfit says, we may bring pains into the nearer future and postpone pleasures. If so, then we may want a shorter duration of pain so we can stop dreading the continuation of the pain at an earlier time. And so, for that reason, we may prefer J to R. But let’s assume that dread is either absent in J-R or counterbalanced by some pleasure. Hence, we may ignore the bias towards the future.

People who have had severe migraines can best judge whether a day of R’s pain is worse than five minutes of J. Most of the migraine sufferers I’ve talked to prefer R. So, I answer “yes” to the first question: migraine pain *can* be so awful that five minutes of it would be at least as bad as a day of the milder pain.

To avoid where this argument is headed, some might object, “R is worse than J even though one should choose R when J is the only alternative. R is worse because it compares less favorably than J to some unavailable possibilities.” This strategy could be deployed against any comparative judgment. But which of the endless unavailable options are relevant to comparing J and R, if comparing them directly doesn’t establish their relative value? And what comparisons could reverse our initial judgment that R is not worse than J? These questions appear unanswerable.^{vi}

Second, can an agonizing J-level headache be transformed into a milder R-level headache with 8 reductions of intensity? Since duration doubles with each move down the alphabet, each change should be for the worse even if the pain intensity is considerably reduced at each step. 8 reductions, I think, is more than we need. Again we should answer “yes.”

Hence, J-R entail that *being intrinsically better than* is not transitive. “More cautiously,” says Temkin, “one may decide that the concept of “better than” is limited in scope, and that for [apparent counterexamples to Transitivity] one needs another concept for

comparing alternatives that is similar in meaning, but intransitive.”^{vii} But one can use an intransitive concept in all of one’s comparative, normative judgments—why multiply concepts?

The Second Counterexample: Long Periods of Pain

The second counterexample consists in the possibilities A-Z, each of which involve a single person’s experience:

A: 1 year of excruciating agony.

B: 100 years of pain slightly less intense than that in A.

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

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

...

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

Z: 10^{50} years of pain slightly less intense than the pain in Y. The pains in Z are slightly worse at each moment than unconsciousness.^{viii}

If you doubt that 25 slight reductions in intensity could turn A’s unpleasures into Z’s, replace “slightly less intense” with “somewhat less intense” or “A-Z” with “1-100.”

This counterexample retains its force even if one amplifies A-Z in various ways. To keep matters simple, one should amplify A-Z only with details that have little or no effect on intrinsic disvalue.

Z is not worse than A; a tremendously long period of Z’s milder pains is not worse than horrible agony for one year. To make this rhetorically compelling, I would now

describe a method of torture that would conjure up horrible agony. But I'll leave that task to your imagination, if you are tempted to think that Z is worse than A.

Although Z is not worse than A, 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 outweighs slightly reducing its intensity. So, the possibilities get worse and worse, but Z is not worse than A. This contradicts Transitivity.

Or does it? According to Transitivity, if A is intrinsically better than B, and B is intrinsically better than C, then A is intrinsically better than C. I interpret "A," "B" and "C" to stand for *any* consistent possibilities, but some philosophers might say, "Evaluative concepts such as *better than* are essentially practical, so we should restrict Transitivity to possibilities that might bear on action." Such a restriction runs counter to the spirit of theoretical philosophy; but anyway, showing that A-Z violates the unrestricted version of Transitivity would, in various ways, support the first counterexample which *does* apply to the restricted version.

Two worthwhile objections don't say how the example goes wrong—just that it must. According to the first, "Sorites arguments—and this is one—are known to be unsound." Sorites arguments appeal to a series of changes that individually make no difference. For example, having one hair fewer makes no difference to whether someone is bald. But in this example, each change matters: each succeeding possibility is worse; each change in intensity makes the pain worse. So, the first counterexample is not of the Sorites type.^{ix}

According to the second undiagnostic objection, "This example can't be assessed; our judgments about such bizarre lives can't be trusted." But each comparison involves just two factors: duration and pain intensity. And most of us have strong beliefs about each comparative judgment; the questions don't strike us as too bizarre or difficult to answer correctly. So, this objection is too quick. But the next objection tries to explain more precisely why our intuitions fail.

Two Diagnostic Objections to the Second Counterexample

(i) “Z is worse than A, although A seems worse because we have difficulty conceiving how much disvalue can accumulate, bit by bit, over 10^{50} years.” But even after reflecting on Z’s length—also bearing in mind that Z’s pain doesn’t worsen over time—I still don’t prefer A to Z. In fact, I strongly prefer Z to A. Should I? The following argument emphasizes the ratio of time spent suffering in A to time spent in the mild pain of Z (where “mild pain” means “pain at each moment slightly worse than temporary unconsciousness”):

1. One million years of mild pain are worse than three seconds of agony; (premise)
2. A period of pain is exactly as bad as the summed badness of its sub-periods of three seconds or more; (premise)
3. One year of agony is exactly as bad as the summed badness of each of its 10 billion three-second sub-periods; (from 2)
4. 10^{16} years of mild pain is exactly as bad as the summed badness of each of its 10 billion million-year sub-periods; (from 2)
5. 10^{16} years of mild pain are worse than one year of agony; (from 1, 3 and 4)
- C. 10^{50} years of mild pain are worse than one year of agony (Z is worse than A). (from 5)

All of the derivations in this argument are valid. So, if the argument is unsound, then either 1 or 2 is false. Therefore, the argument proves that the following triad is inconsistent:

1. One million years of mild pain are worse than three seconds of agony.
2. A period of pain is exactly as bad as the summed badness of its sub-periods of three seconds or more.

~C. 10^{50} years of mild pain are not worse than one year of agony (Z is not worse than A).

At least one of these is false. But which? Any member of the triad can be denied by appealing to the other two, so let's look for independent evidence.

Thoughtful people who know suffering overwhelmingly prefer an indefinitely long period of Z's pain to a year of agony. So, ~C is well-supported independently of the trilemma. (In fact, a slightly stronger thesis is also well-supported: that Z is preferable to A.)

According to Moore, "two bad things . . . may form a whole much worse than the sum of badness of its parts."^x Is a period of pain worse than the summed disvalue of its temporal parts? A period of pain might be worse than the summed disvalue of its parts that last only a millionth of a second; such parts may be too brief to have disvalue. But 2 says, "A period of pain is exactly as bad as the summed badness of its sub-periods *of three seconds or more*," and one can suffer in three seconds. Is 2 true? 2 seems true for the following reason: a period of pain is bad because of how it feels or what it's like; and how it feels or what it's like consists in how its subperiods feel or what they're like. 2 does presuppose the controversial thesis that badness can be summed, but we can revise 2 as follows:

2R: The reasons why a period of pain is bad are exhausted by the reasons why each of its subperiods of at least three seconds are bad.

2R can replace 2 in our discussion. And the same reasoning that supports 2 supports 2R. Hence, 2R is well-supported independently of the trilemma.

Now consider 1: "Three seconds of agony are preferable to a million years of mild pain." Is this true of the worst agony? I find the issue too hard to decide intuitively. Many other people I know agree.^{xi}

According to Gurney, “torture” is “incommensurable with moderate pain”—so, any duration of torture is worse than any duration of moderate pain.^{xii} Perhaps “moderate pain” is roughly at the level of Z’s pains. If so, Gurney rejects 1. Sidgwick responds as follows:

[Gurney’s] doctrine . . . does not correspond to my own experience; nor does it appear to me to be supported by the common sense of mankind:—at least I do not find, in the practical forethought of persons noted for caution, any recognition of the danger of agony such that, in order to avoid the smallest extra risk of it, the greatest conceivable amount of moderate pain should reasonably be incurred.^{xiii}

Sidgwick’s reply assumes that, if an indefinitely long period of Z’s pain is preferable to a short duration of agony, then one should choose that long period of pain over a *slight* chance of the agony. Is this true? Let “F” represent behavior that has a 100% chance of causing an indefinitely long period of Z’s pain, while “G” represents behavior that has a .1% chance of causing 3 seconds of horrible pain followed by normal life and a 99.9% chance of being followed by normal life. Is F wiser than G, on Gurney’s view? Perhaps not: someone who has a “normal life” might occasionally enjoy ecstasy, which might counterbalance the possibility of three seconds of agony. But even if F were wiser than G, even cautious persons would probably opt for G. None of us want certain pain, and we’re all prone to disregard small chances.

Hence, no solid, independent evidence counts for or against 1. And 1 conflicts with the conjunction of $\sim C$ and 2R, which independent evidence supports. So, 1 is false; Z is not worse than A.

(ii) “At least one of our other comparative judgments is false; some period of slightly more intense pain isn’t worse than a period of slightly less intense pain that lasts 100 times longer.” This thesis is tempting only when the lesser pain is mild: perhaps a period of pain isn’t worse than a period of slightly lesser, mild pain that lasts 100 times longer. However,

no one believes that a severe pain's becoming slightly less intense outweighs its duration increasing 100-fold. So, perhaps W is worse than V, but A is certainly not worse than B.

This objection, though tempting, is utterly implausible given that even Z's pains are worse than temporary unconsciousness. If pain is that bad or worse, then reducing its duration by 99% would obviously be worth a slight increase in intensity. So, Transitivity fails.

Again, note that one can easily construct a similar example involving pleasure. A is a year of ecstasy, B is 100 years of pleasure slightly less intense, C is 10,000 years of pleasure still slightly less intense, and so on. Z is 10⁵⁰ years of experience slightly better than temporary unconsciousness at each moment—the pleasures of muzak and potatoes.^{xiv} B is better than A, C is better than B, and so on, but Z is not better than A. This contradicts Transitivity.

Rational Choice

Transitivity's failure entails that, for some possibilities, X is better than Y, Y is better than Z, but X is not better than Z. To simplify the coming discussion, I'll adopt the stronger thesis (which I believe) that for some possibilities, X is better than Y, Y is better than Z, and X is *worse* than Z.^{xv} This thesis may seem to run afoul of the following "money-pump" objection^{xvi} (I am not quoting):

On your view, an informed, rational person may prefer X to Y, Y to Z, and Z to X. But such preferences lead to irrational behavior. For example, in some circumstances, she would pay a small amount to trade X for Z, then pay a small amount to trade Z for Y, then pay a small amount to trade Y for X—the same X she started with. Then she would lose more money in the same way, again trading X for Z, Z for Y and Y for X.^{xvii}

But, on my view, an informed, rational person won't prefer X to Y, Y to Z, and Z to X in a sense that commits her to such insanity.^{xviii} She would prefer X to Y were they alone relevant to which of them she should prefer—X is, after all, better than Y. And, similarly, she would prefer Y to Z and Z to X. But when all three possibilities are available, she won't prefer X to Y, Y to Z, and Z to X, to avoid being money-pumped. Hence, a rational person will not always prefer what is better. Given a choice among X, Y and Z, either Z's presence bears on whether she should prefer X to Y, or X's presence bears on whether she should prefer Y to Z, or Y's presence bears on whether she should prefer Z to X. For example, starting off with X, she may consistently (and wisely) refuse to trade it for Z, even knowing that Z is better. Unusual cases often need unusual treatment.

But what should she do: stay with X; trade X for Z and then stay put; or trade X for Z, then trade Z for Y and then stay put?

For cases that violate Transitivity, one lacks a powerful reason to justify any particular course of action: one cannot say *this choice results in a state of affairs that is best*. But there may be reasons, regarding the possibilities themselves, to prefer some choices over others. For example, perhaps we have special reason to avoid a possibility with *much* greater pains than occur in some other option. (So, in the second counterexample, we have special reason to avoid A, given that Z is available.) If such principles entail that one of X, Y and Z is most rationally targeted, then an informed, rational person could be money-pumped only until she lands on that choice. If, alternatively, there are no reasons, regarding the possibilities themselves, to prefer some choices over others, then all available choices are equally wise (which, of course, doesn't entail that all possibilities are equally good). If so, then she could not be pumped of a cent: she would keep what she originally has (whether X, Y or Z), rather than giving up a quarter to trade for something better.

Transitivity may persist in our reasoning as a rule of thumb, as exceptions to it are rare. Avoiding the principle altogether would require making 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 would need to compare A and C directly.

So, I see no reason to accept Parfit's view, reported by Temkin, that Transitivity's failure would entail skepticism about practical reasoning.^{xix}

Conclusion

According to Transitivity, if A is intrinsically better than B, and B is intrinsically better than C, then A is intrinsically better than C. I offered general reasons, in the first section, for thinking that Transitivity is not analytically true. Transitivity's support, instead, is merely inductive. But if so, then a sufficiently compelling counterexample can overturn it.

I offered two counterexamples to Transitivity. The first consists in nine headaches, ranging from five minutes to a day. This example, I think, outweighs Transitivity's inductive support all by itself. And it may constitute my best case against Transitivity as a conceptual truth; false propositions aren't conceptually true.

The second counterexample consists in 26 lives, ranging from a year to 10^{50} years. On one objection to it, Z is worse than A: 10^{50} years of mild pain are worse than a year of agony. I discussed this objection in terms of the following inconsistent triad:

1. One million years of mild pain are worse than three seconds of agony.
 2. A period of pain is exactly as bad as the summed badness of its sub-periods of three seconds or more.
- ~C. 10^{50} years of mild pain are not worse than one year of agony.

On this objection, 1 and 2 are true. I accept 2 in revised form. But I reject 1 and accept ~C, that one year of agony is at least as bad as 10^{50} years of pain slightly worse than

unconsciousness. ~C seems obviously true to many of us when we reflect on our worst pains. The best response to the second counterexample, in my opinion, is the following:

The fact that ~C seems to be true counts only minimally in its favor, since we have trouble imagining how much badness accumulates over Z's 10^{50} years. Such modest evidence is less than the inductive evidence for Transitivity. The second example, therefore, fails on its own to refute Transitivity.

I think that reflection supports ~C more than minimally, but I have no argument to offer. Because this issue is so difficult, the first counterexample is better than the second. But they work together against Transitivity. Together they outweigh the inductive evidence that supports that principle. "Being intrinsically better than," I conclude, is nontransitive, as is "being hedonically better than."

Since Transitivity fails, so does the linear picture of value: possibilities cannot all be represented as points on a line, with the intrinsically better ones represented to the right of those intrinsically worse; no complete ordinal ranking of states of affairs exists. If so, then, of course, no complete cardinal ranking exists; possibilities can't all be quantified so that the better ones have higher numbers than those worse (never mind other possible constraints on cardinality, for example, that equal numerical differences always correspond to equal differences in value).^{xx} Moreover, my counterexamples show that the hedonic value of possibilities can't be similarly conceived, even though philosophers often think of hedonic value as a paradigmatic (or as the sole) type of value that can be precisely quantified.^{xxi} If hedonic value escapes neat and comprehensively quantification, this increases the chances that other spheres of value do as well (provided there are others).

Philosophers often discuss the ontology of value rather dismissively, bringing it up only to deny that "value" has a referent.^{xxii} But most of us, prior to argument, are firmly wedded to Transitivity; moreover, quantitative expressions pervade ethical theory, for example, "average utility" and "the principle's weight." Even "more value" and "less value"

connote quantity. This suggests that, at some level, philosophers take the linear view of value seriously—not just as a convenient heuristic, but as representative of value itself. For that reason, it may be worthwhile to note that Transitivity’s failure entails that value itself (if such an item exists) is neither precisely quantifiable nor robustly isomorphic with items that are.

Stuart Rachels

ⁱ John Broome argues for Transitivity in *Weighing Goods*, pp. 10-12.

ⁱⁱ See Larry S. Temkin, “Weighing Goods: Some Questions and Comments,” pp. 361-363 and “A Continuum Argument for Intransitivity,” pp. 193-194. Temkin formulates Transitivity in terms of “all things considered better than” instead of “intrinsically better than,” but this doesn’t matter, since he discusses only properties intrinsic to A, B and C.

ⁱⁱⁱ See Larry S. Temkin, “Weighing Goods: Some Questions and Comments,” p. 363, and especially “A Continuum Argument for Intransitivity,” sect. 4.

^{iv} Thomas Kuhn: “all revolutions involve, among other things, the abandonment of generalizations the force of which had previously been in some part that of tautologies. Did Einstein show that simultaneity was relative or did he alter the notion of simultaneity itself? Were those who heard paradox in the phrase ‘relativity of simultaneity’ simply wrong?” (*The Structure of Scientific Revolutions, second edition*, pp. 183-184)

^v Derek Parfit, *Reasons and Persons*, p. 160.

^{vi} See Larry S. Temkin, “Intransitivity and the Mere Addition Paradox,” pp. 180-183 and his discussion of the same material in “Rethinking the Good, Moral Ideals and the Nature of Practical Reasoning,” sect. L.

^{vii} Larry S. Temkin, “Weighing Goods: Some Questions and Comments,” p. 361, fn. 22.

^{viii} What happens to these poor suffering souls after the year in A, the century in B, and so on? Three variations are death, normal life and Heaven (God’s apology). But what happens *later* doesn’t affect the *intrinsic* disvalue of these possibilities.

^{ix} My response follows Derek Parfit, “Overpopulation and the Quality of Life,” p. 160, fn. 12. Temkin (‘A Continuum Argument for Intransitivity,’ sect. 5) rebuts this objection differently.

^x G. E. Moore, *Principia Ethica*, p. 28.

^{xi} “I’ll let the theory take care of that one.”—John O’Leary-Hawthorne.

^{xii} Edmund Gurney, *Tertium Quid*, Vol. I, p. 181. Compare Dostoevsky on pleasure: “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” (from Geir Kjetsaa, *Fyodor Dostoevsky: A Writer’s Life*, p. 149).

^{xiii} Henry Sidgwick, *The Methods of Ethics*, pp. 123-124, fn. 1.

^{xiv} Parfit prefers “the Century of Ecstasy” to “the Drab Eternity” of muzak and potatoes, and Griffin concurs. If so, then Parfit and Griffin would prefer A to Z. See Parfit, “Overpopulation and the Quality of Life,” pp. 160-161. (This repeats his view in *Reasons and Persons*, pp. 498-499.) And see Griffin, *Well-Being*, p. 86.

^{xv} I’ll state dogmatically why I believe this. There should be possibilities involving pain such that Y’s greater duration outweighs X’s greater intensity, Z’s greater duration outweighs Y’s greater intensity, but X’s greater intensity outweighs—and not merely balances—Z’s greater duration. Let me try to identify such possibilities in the first counterexample, which consists in the J-R headaches. R is a day of pain slightly worse than unconsciousness. R is preferable to J, which is five minutes of agony. I want the possibility earliest in the alphabet that is preferable to J, J, and an intermediary possibility. I’m not sure which is first preferable to J, but suppose P is. P is the six hour headache. Now consider J, P and the possibility equidistant from them:

J = five minutes of agony

M = forty minutes of a headache that is bad but not nearly as bad as J’s

P = six hours of a headache that is bad but not nearly as bad as M’s.

M’s duration outweighs J’s intensity; P’s duration outweighs M’s intensity; but J’s intensity outweighs P’s duration. So, M is worse than J; P is worse than M; but J is worse than P.

^{xvi} The first such argument, to my knowledge, appeared in Donald Davidson, J. C. C. McKinsey and Patrick Suppes, “Outlines of a Formal Theory of Value, I,” p. 146. The authors say, “We owe the inspiration for this example to Dr. Norman Dalkey of the Rand Corporation.”

^{xvii} In my discussion, X, Y and Z are possibilities. One cannot trade possibilities, but one can trade the means to making them obtain. I ignore this wrinkle in the text.

^{xviii} Compare Robert Nozick, *The Nature of Rationality*, p. 140, fn.

^{xix} As reported by Temkin, “A Continuum Argument for Intransitivity,” p. 209.

^{xx} See James Griffin, *Well-Being*, ch. VI for a nice discussion of measuring value.

^{xxi} See Plato, *Protagoras*, 356b and 357 a-e; John Stuart Mill, *Utilitarianism*, ch. V, fn., three paragraphs before the end; Richard Brandt, *A Theory of the Good and the Right*, p. 255; James Griffin, *Well-Being*, pp. 75, 104. For three related, relevant arguments, see Henry Sidgwick, *The Methods of Ethics*, p. 94; G. E. Moore, *Principia Ethica*, p. 78; David Wiggins, “Weakness of Will, Commensurability, and the Objects of Deliberation and Desire,” p. 267 (p. 255 as reprinted in *Essays on Aristotle’s Ethics*).

^{xxii} Nagel says, in *The View From Nowhere*: “What we aim to discover is not a new aspect of the external world, called value, but rather just the truth about what we and others should do and want.” (p. 139) “The objective badness of pain, for example, is not some mysterious further property that all pains have, but just the

fact that there is a reason for anyone capable of viewing the world objectively to want it to stop.” (p. 144)

Korsgaard says, “To talk about values . . . is not to talk about entities, either mental or Platonic, but to talk in a shorthand way about relations we have with ourselves and one another.” (*The sources of normativity*, p. 138)