

# Better than Nothing

Campbell Brown

Draft revised 27-04-2012

## Abstract

A good life, or a life worth living, is a one that is “better than nothing”. At least that is a common thought. But it is puzzling. What does “nothing” mean here? It cannot be a quantifier in the familiar sense, yet nor, it seems, can it be a referring term. To what could it refer? This paper aims to resolve the puzzle by examining a number of analyses of the concept of a life worth living. Temporal analyses, which exploit the temporal structure of lives, are distinguished from non-temporal ones. It is argued that the temporal analyses are better.

## Introduction

The concept of *a life worth living* is prominent in what is sometimes known as “population ethics”, the study of ethical problems involving how many people there are, and who they are. It features, for example, in Derek Parfit’s famous “Repugnant Conclusion”:

For any possible population of at least ten billion people, all with a very high quality of life, there must be some much larger imaginable population whose existence, if other things are equal, would be better, even though its members have *lives that are barely worth living*. [10, 388, italics added by me]

Parfit is not inclined to accept to this conclusion (hence the name he gave it). His problem, however, is finding a plausible theory that will allow him consistently to reject it. My goal here is not to propose a solution, but rather something more modest. If we want to know why the Repugnant Conclusion is false — or indeed *whether* it is false — then a good start would be to know what it *means*. And that means understanding what it is for a life to be worth living.<sup>1</sup> My aim, then, is to examine various analyses of this concept, to see whether any is satisfactory. If one can grasp a concept without an analysis, it seems nonetheless reasonable to hope that reflecting on possible analyses might at least strengthen one’s grasp.

In population ethics, especially, the concept of a life worth living may be put to useful service fixing a zero-point in the scale of well-being by reference to which comparisons of total well-being may then be made. One plausible criterion of goodness for possible outcomes, or worlds, is total well-being.<sup>2</sup> That the people in one outcome enjoy greater

---

<sup>1</sup>I assume “barely worth living” entails “worth living”, so understanding the former entails understanding the latter.

<sup>2</sup>By a person’s “well-being” I mean how good the person’s life is. A statement like “A’s well-being is greater than B’s” is equivalent, on my usage, to “A’s life is better than B’s”.

total well-being than those in another is a reason to prefer the former; other things being equal, the greater the total well-being, the better. In the context of population ethics, however, where population size may vary, this criterion presupposes that we know what zero well-being is.

To illustrate, suppose we measure the well-being of every person in each of two outcomes,  $X$  and  $Y$ . We then add up our measurements and determine, say, that  $X$  has the greater total. However, just when we're about to declare  $X$  the winner (we've checked that in all other respects the outcomes are equal), we realise that our measuring device had not been properly calibrated, so that the zero-point was wrongly set. To compensate for this, we uniformly add a constant,  $\beta$ , to all our measurements, and then add up again. What effect, if any, might this correction have on our resulting judgement? What we've done is equivalent to adding  $m\beta$  and  $n\beta$  to our total measurements for  $X$  and  $Y$ , respectively (where  $m$  and  $n$  are the numbers of people, or population sizes, in  $X$  and  $Y$ , respectively). If  $m = n$ , i.e. population size is constant, then, obviously, this will make no difference to which total is greater; we needn't have bothered with the correction. However, if  $m \neq n$ , then the correction could tip the scales: if the difference  $n\beta - m\beta$  is at least as great as the difference between our uncorrected total measurements, then our new result will be that  $X$  does not have greater total well-being, after all. What this tells us is that, insofar as we care about total well-being, the zero-point matters when population size varies (and only then).

How then are we to find the correct zero-point? This is where the concept of a life worth living may earn its keep. A common approach is to define *zero well-being* as the boundary between lives which are worth living, above zero, and lives which are “unworth” living, below. This is what is going on in the Repugnant Conclusion: because the people in the second, larger population have lives “barely worth living”, their well-being, though small, is nonetheless *positive*, and so provided there are enough of them, the total well-being in this population will exceed that in the first, smaller population, even though in the smaller population people have a “high quality of life”. One lesson to have been drawn from this is that total well-being cannot be all that matters; we should reject what Parfit calls the “Impartial Total Principle” [10, p. 387].

I take as my starting point the common thought expressed as follows by Parfit:

[A] life of a certain kind may be judged to be either good or bad — either worth living, or worth not living. If a certain kind of life is good, it is better than nothing. If it is bad, it is worse than nothing. [10, p. 487]

Though common, this view is, on reflection, puzzling. What does it mean to say of something that it is “better than nothing”? What does “nothing” mean here? There are two obvious interpretations, neither satisfactory.

A similar ambiguity is exploited to amusing effect by Lewis Carroll:

“Who did you pass on the road?” the King went on . . .

“Nobody,” said the Messenger.

“Quite right,” said the King: “this young lady saw him too. So of course Nobody walks slower than you.”

“I do my best,” the Messenger said in a sulky tone. “I'm sure nobody walks much faster than I do!”

“He can’t do that,” said the King, “or else he’d have been here first. . . .” [7, p. 197]

The joke here is that “nobody” is treated as a *quantifier* by the Messenger, but as a *name* by the King. (The messenger’s interpretation seems the more orthodox.) Similarly, the “nothing” in “better than nothing” could be treated either as quantifier or name.

On the quantifier interpretation, “ $x$  is better than nothing” is equivalent to “for all things  $y$ , it is not the case that  $x$  is better than  $y$ ”. This cannot be what Parfit intends. As “faster than nobody” would mean (one of) the slowest, so “better than nothing” would mean (one of) the worst. But Parfit surely does not mean to say that a life worth living is the *worst* sort of life. Bad lives must be worse than good ones.

Perhaps, then, “nothing” is a name. But the name of what? In mathematics, “nothing” might sometimes refer to the empty set, i.e. the set with no members. Similarly, in logic it might refer to a tautology (because, e.g., asserting a tautology is like asserting nothing, one might say). I have nothing against these ways of speaking in these contexts, but I find it unhelpful here. I can make sense of comparisons of value between lives. I have an idea of what it means to say, for example, that one life is better than another. But the empty set is not a life, so far as I can tell, and I’m lost when it comes to comparing a life, in this way, with a non-life such as the empty set. Is such-and-such a life better than the empty set? I don’t understand the question.

There appears not to be a third interpretation, at least not one as obvious and simple as the first two. So we’re left with a puzzle. Parfit seems aware of this. He continues:

Judgements of this kind [i.e. of being better or worse than nothing] are often made about the last part of some life. Consider someone dying painfully, who has already made his farewells. This person may decide that lingering on would be worse than dying. . . . And he might in a similar way decide that he was glad about or regretted what lay behind him. He might decide that, at some point in the past, if he had known what lay before him, he would or would not have wanted to live the rest of his life. He might thus conclude that these parts of his life were better or worse than nothing. If such claims can apply to parts of a life, they can apply, I believe, to whole lives. [10, p. 487]

Parfit is on the right track here, I think. His comments are in some ways suggestive of the Limit Analysis, which I will endorse below. But this is not yet a solution to the puzzle. It is one thing to profess a belief that something is possible, another to explain how. The latter is what the analyses examined below attempt to do.

My goal, I should emphasise, is to find an analysis that is neutral between rival substantive theories of the value of lives, or theories of well-being, as they are commonly known. The philosophical literature provides a number of such theories. For example, there is *hedonism*, according to which the more pleasure a life contains the better it is, and the more pain the worse, other things equal. And there is *desire-satisfactionism*, according to which a life is better the more similar it is to what some appropriately situated agent would desire it to be. And so on. The analysis I seek is not tied to any one of these theories.

Consider, for instance, this proposed analysis: a life is worth living iff it contains more pleasure than pain. This may be fine for hedonists (though I shall argue below, Section 1.1, that it has problems even there), but it is clearly no good for non-hedonists. Take a pluralistic theory according to which there are goods in addition to pleasure, say, knowledge, and bads in addition to pain, say, ignorance. Someone who held this theory

may believe that a life containing more pain than pleasure could nonetheless be worth living if it also contained enough knowledge. The proposed analysis is, therefore, incompatible with pluralism of this kind. And this is a reason to reject the analysis. The issues dividing hedonists and pluralists — e.g., whether knowledge has irreducible value — are substantive moral issues. They are to be settled by the usual methods of moral theory, not stipulated out of existence by definitional fiat. (This is not to say, of course, that the proposed analysis is *false*, but only that it is not acceptable as an analysis.)

Perhaps it will seem that I am setting my sights too high. Hoping to discover an analysis compatible with with *every* conceivable theory of well-being might seem overly optimistic. And indeed, as we will see, such scepticism is largely borne out by the investigation to follow: for each analysis canvassed, we will see incompatible theories. But even if all analyses fall short of the ideal, some may fall less short than others. If the theories excluded by an analysis are independently plausible and well-established in the literature, then excluding them would be a more significant cost than excluding theories which, though internally coherent, have little else going for them. We should be less concerned about blocking off regions of logical space that no one has wanted to occupy or could have much reason for occupying.

The analyses examined below are divided into two categories. Lives, I assume, are temporally structured; they change over time. What I call “temporal analyses” exploit this structure, whereas “non-temporal analyses” do not. I begin with the non-temporal (Section 1), of which I consider three: the “Balance Analysis”, which says a life is worth living iff it contains more good than bad (Section 1.1 and Section 1.2); the “Empty Life Analysis”, which says a life is worth living iff it is better than a life containing no good and no bad (Section 1.3); and the “Worlds Analysis”, which says a life is worth living iff it is better for a person to live this life than not to exist (Section 1.4). All of these, I argue, have significant problems.

I then turn to temporal analyses (Section 2). First, I introduce a simple mathematical model incorporating the dimension of time (Section 2.1). I then discuss two temporal analyses: the “Flatline Analysis”, which says a life is worth living iff it is better than a life the value of which over time is constant (Section 2.2); and the “Limit Analysis”, which says a life is worth living iff the life as a whole is better than an “infinitesimal truncation” of the life (Section 2.3). I argue that these are superior to the non-temporal analyses, because they largely avoid the problems identified previously (Section 2.4). Between the two temporal analyses, I find it hard to choose. I argue that each has a slight advantage over the other (Section 2.5).

Before continuing, two brief notes on terminology. First, in what follows, for the sake of brevity, I’ll often say simply “a good life”, rather than “a life worth living”; these should be taken as synonymous. Second, if a good life is one that is *better* than nothing, then there are two ways for a life to fail to be good: it may be either *worse* than nothing, or *equally as good as* nothing. In the former case I’ll say the life is “bad”, and in the latter that it is “neutral”. When I discuss each of the various analyses below, I will explicitly state only the proposed definition of a good life. The corresponding definitions of bad and neutral lives should be obvious. Roughly, if the definition stated has the form “*L* is good iff *X* is greater than *Y*”, then the corresponding definitions would be “*L* is bad iff *X* is less than *Y*” and “*L* is neutral iff *X* is equal to *Y*”.

## 1 Non-temporal Analyses

### 1.1 The Balance Analysis I

Most lives are a mixture of good and bad. They are in some ways good, in other ways bad. Some aspects of the life make it better: without them, it would be worse than it is. Other aspects make it worse. A common approach to judging a life good or bad involves weighing up, or *balancing*, the good and bad elements of the life. When the good outweighs the bad, the life as a whole is judged good, and so on. On this approach, it is natural to associate “nothing” with a balance of *zero*, i.e., with the balance of a life that contains exactly equal, perfectly balanced, quantities of good and bad stuff. A life may then be regarded as better than nothing, and hence good, just in case it has a positive (greater than zero) balance, a surplus of good over bad.

Perhaps an apt analogy would be with a bank account. The balance of an account is the sum of all credits to the account minus the sum of all debits. Suppose you discover that your account’s balance is less than you hoped it would be, yet still positive. You might say, tempering your disappointment, “Well, at least that’s better than nothing”. The balancing approach suggested above thinks of the value of a life on the model of a bank account. When something good happens in a person’s life — say, she experiences some pleasure — an amount of well-being is credited to her “well-being account”. When something bad happens, a corresponding amount is debited. The value of her life is given by the balance of her well-being account. Her life is good just in case her account is overall “in credit”, i.e., total credits exceed total debits.

This approach yields the following analysis.

*The Balance Analysis.* A life is good iff the amount of good stuff it contains exceeds the amount of bad.

By “good stuff” I mean whatever makes a life better by its presence: the more of this stuff a life contains, the better it is, other things being equal [Cf. 2, p. 169]. By “bad stuff” I mean whatever makes a life worse, less worth living. For example, hedonism, as discussed earlier, says the only good stuff is pleasure, and the only bad pain.

One limitation of this analysis is that it presupposes the existence of good and bad stuff. Some theories of well-being are not aptly characterised in such terms; they are a poor fit for the bank account model. Consider, e.g., a preference-based theory: one life is better for you than another just in case you prefer it. Such a theory says nothing of good and bad stuff. Moreover, your preferences may be too “holistic” to be reduced simply to a preference for more of one sort of stuff and less of another. Yet surely this theory should not for this reason be debarred from classifying lives as good or bad.

Moreover, even friendly theories like hedonism are problematic. Within the hedonistic camp, we can find theories of well-being that are incompatible with the Balance Analysis. To show this, it will be useful to introduce a simple formal model. Let a *life* be an ordered pair  $(x, y)$ , where  $x$  and  $y$  are non-negative real numbers representing, respectively, the total quantities of pleasure and pain contained in the life. We need to assume that these quantities are somehow measurable on a common scale, otherwise saying that a life contains more pleasure than pain would be as meaningful as saying, e.g., that today I drank more coffee than I watched television. One might doubt this assumption. Were quantity merely duration, then comparing quantities of pleasure and pain might be unproblematic. We could, in principle at least, add up the total time in a life during which pleasure

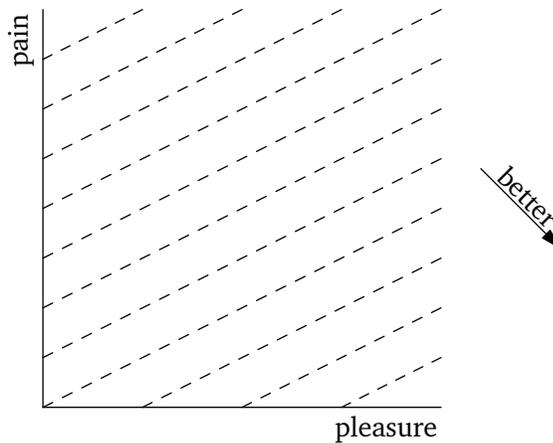


Figure 1: first theory of well-being

is experienced, and compare this with the total time during which pain is experienced. Presumably, however, quantities of pleasure and pain depend not only on duration, but also on intensity. Perhaps five minutes of strong pleasure is a greater quantity of pleasure than, say, fifteen of mild. One might doubt, however, whether intensity of pleasure can meaningfully be compared with intensity of pain. If so, this would be more bad news for the Balance Analysis. But, for argument's sake, I shall grant the assumption.

I assume, further, that the value of lives may be represented by a value function  $v$ , on a cardinal scale.<sup>3</sup> Since the representation is merely cardinal, both the *unit* and the *zero* are arbitrary. The latter is especially important in the present context. It is common in population ethics to assume a numerical representation whereby zero is understood as standing for neutrality. We cannot, however, follow this tradition here. It would be circular to define  $v$  in terms of the very thing we wish to analyse with it. Once we've found a satisfactory analysis, we can then use it to transform  $v$ , if necessary, so zero represents neutrality. But we cannot simply assume that it has this property at the outset.

Now, within the hedonist camp, there are various views about the relative importance of pleasure and pain. Some hold, for example, that achieving pleasure is less important than avoiding pain. Forgoing pleasure in order to avoid pain (e.g., drinking in moderation to avert a hangover) may be rational even in some cases where the pleasure forgone is *greater* than the pain avoided. As this view is sometimes put, pleasure is less good than pain is bad.<sup>4</sup> To take a concrete example of this view, let us suppose the following value function:  $v(x, y) = x - 2y$ . On this view, we might say, pleasure is only half as good as pain is bad: the amount by which a life would improve due to the addition of some quantity of pleasure is only half the amount by which it would improve due to the subtraction of an equal quantity of pain.<sup>5</sup> This theory is shown in Figure 1. The dashed lines are indifference

<sup>3</sup> Here's what I mean by this. The function  $v$  is an arbitrarily selected member of a set of functions  $X$  such that  $X$ , and no non-empty proper subset of  $X$ , is closed under positive affine transformations. A positive affine transformation of  $v$  is a function  $v'$  such that, for some numbers  $\alpha$  and  $\beta$  (with  $\alpha > 0$ ),  $v'(l) = \alpha v(l) + \beta$  for every  $l$ , i.e.  $v'$  results from  $v$  by changing at most the unit or zero. Thus only those properties of  $v$  which are preserved under all such changes may be taken to have any significance. In particular, the order of the numbers  $v$  assigns to lives and the ratios of differences between those numbers may be significant; but, e.g., whether the numbers are positive or negative cannot be.

<sup>4</sup>See e.g. 8.

<sup>5</sup>Formally, for any positive  $\beta$ ,  $v(x + \beta, y) - v(x, y) = \frac{1}{2}(v(x, y - \beta) - v(x, y))$ .

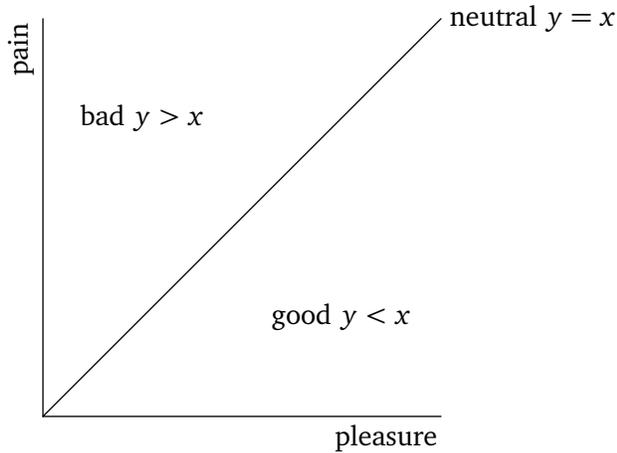


Figure 2: The Balance Analysis

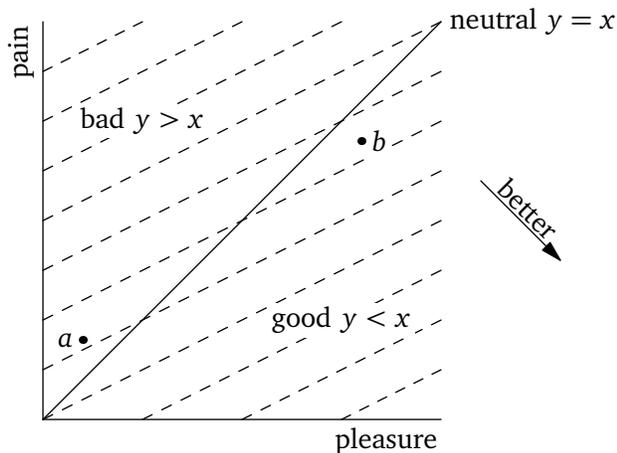


Figure 3: first theory of well-being combined with the Balance Analysis

curves, i.e., equivalence classes under the relation “equally good”. Any two points on the same line represent lives that are equally good. The direction of “betterness” is, as shown, towards the south-east: lives get better as we move either east, increasing pleasure, or south, reducing pain.

The Balance Analysis may be depicted as in Figure 2. The line at  $45^\circ$  represents those lives in which the balance of pleasure over pain is zero ( $y = x$ ); these lives are neutral, according to the Balance Analysis. Lives to the south-east (north-west) have a positive (negative) balance, and so are good (bad).

To see the disastrous consequences of combining the Balance Analysis with the theory of well-being suggested above, simply superimpose the previous two graphs, as in Figure 3. Notice, the neutral line intersects more than one indifference curve. This allows us to find lives such as those represented by points  $a$  and  $b$  on the graph. Because  $a$  is below and  $b$  above the same indifference curve,  $a$  is better than  $b$ . But because  $a$  is above and  $b$  below the neutral line,  $a$  is bad and  $b$  good. We have thus reached the absurd conclusion that some good lives are worse than a some bad ones.

Not only is this highly counter-intuitive, but it defeats one of the central purposes for

which we were interested in the concept of good life in the first place. As explained above, the idea is to use this concept to fix a zero-point in our utility scale, so that good lives receive positive utilities, and bad lives negative. But this won't work if some bad life is better than some good life, because a better life must receive a greater utility than a worse one.

To emphasise, I am not claiming that this theory of well-being is true. I claim only that it is not so obviously implausible that we should be happy to rule it out by our choice of an analysis of good and bad lives. An analysis that is compatible with this theory would be preferable, other things equal.

## 1.2 The Balance Analysis II

Perhaps it will be objected that this argument rests on an uncharitable interpretation of the Balance Analysis. To say that the good stuff in a life outweighs the bad is not, according to this objection, to say simply that there is more good stuff than bad; it is to say, rather, that the *goodness* of the good stuff is greater than the *badness* of the bad. What needs to be balanced is not sheer quantities of good and bad stuff, but rather “quality-adjusted” quantities, as we might say. Consider the life (3, 2). If pain is twice as bad as pleasure is good, then two units of pain is more bad than three units of pleasure is good, and so, even though this life contains more pleasure than pain, the goodness of the pleasure is *less* than the badness of the pain. The Balance Analysis properly interpreted, therefore, classifies this life as bad, not good (given the theory of well-being we have been assuming).

Thus, we have another version of the Balance Analysis:

*The Balance Analysis (Quality-adjusted).* A life is good iff the goodness of the good stuff it contains exceeds the badness of the bad.

What is meant here by “the goodness of the good stuff”? It is natural to think of this in terms of the marginal contribution that the good stuff makes to the value of the life. Earlier I suggested that what makes good stuff good is its positive marginal contribution: removing it would worsen the life, other things equal. The *extent* of its goodness is then naturally understood as the magnitude of its marginal contribution, the amount by which it makes the life better.<sup>6</sup> That is, the goodness of the good stuff in a life equals the amount by which the value of the life would drop were the quantity of good stuff reduced to zero, while holding the quantity of bad stuff fixed. Likewise, on this approach, the badness of pain in a life equals the magnitude of the (negative) marginal contribution it makes to the value of the life. Formally, then, the goodness of pleasure and badness of pain in a life  $(x, y)$  may be represented by the functions  $g(x, y) = |v(x, y) - v(0, y)|$  and  $b(x, y) = |v(x, y) - v(x, 0)|$ , respectively. The Balance Analysis may then be interpreted as saying that a life  $(x, y)$  is good iff  $g(x, y) > b(x, y)$ .

This works out nicely if the value function  $v$  satisfies three conditions. First, say that  $v$  is *pleasure-friendly* iff  $v(x, y) > v(x', y')$  whenever  $x > x'$  and  $y \leq y'$  (a life with more pleasure and no more pain is better). Second, say that  $v$  is *pain-averse* iff  $v(x, y) > v(x', y')$  whenever  $x \geq x'$  and  $y < y'$  (a life with less pain and no less pleasure is better). Finally, say  $v$  is *additively separable* iff there exist functions  $f_1$  and  $f_2$  such that  $v(x, y) = f_1(x) + f_2(y)$ . It follows that, if  $v$  is pleasure-friendly, pain-averse, and additively separable, then

---

<sup>6</sup>This assumes that  $v$  represents the value of lives on a cardinal scale.

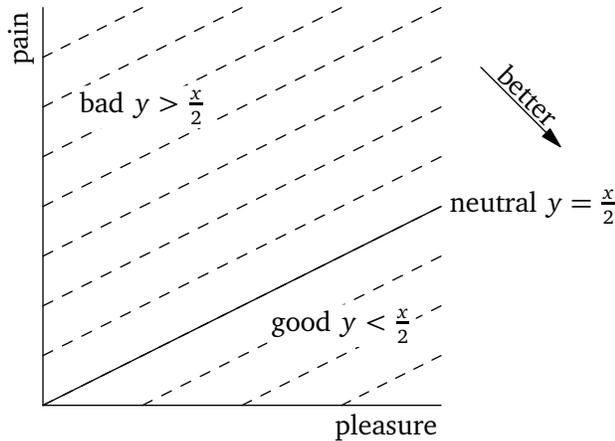


Figure 4: first theory of well-being combined with revised Balance Analysis

$g(x, y) - b(x, y) = v(x, y) - k$ , where  $k$  is the constant  $f_1(0) + f_2(0)$ .<sup>7</sup> So, according to the Balance Analysis,  $(x, y)$  is good only if  $v(x, y) > k$ , and bad only if  $v(x, y) < k$ ; so a good life must be better than a bad life, as desired.

The value function considered above,  $v(x, y) = x - 2y$ , satisfies the three conditions. Combined with the revised Balance Analysis, this function implies that  $(x, y)$  is good iff  $y < x/2$ , as shown in Figure 4. Notice in this case the neutral line intersects exactly one indifference curve; indeed the neutral line *is* one of the indifference curves. Clearly, then, all lives above this line are worse than all below.

Things are not so rosy, however, for value functions that are not additively separable. Here's an example:

$$v(x, y) = \frac{x + 1}{y + 1}.$$

Indifference curves for this value function are shown in Figure 5. Given this function,  $g(x, y) = b(x, y)$  iff  $y = x/(x + 1)$ , as represented by the neutral curve in the graph. This curve intersects more than one indifference curve, and so again we can find lives  $a$  and  $b$ , as shown, such that  $a$  is bad,  $b$  is good, and  $a$  is better than  $b$ .

Again, I suggest, the theory of well-being represented by this value function is not obviously crazy. Notice, for one thing, the value function is both pleasure-friendly and pain-averse. We should, therefore, try not to exclude it. We should resist building additive separability into our analysis of the concept of a good life.

### 1.3 The Empty Life Analysis

Consider another analogy. You're sharing a jug of beer with a group of friends. You go around the group pouring some beer into each person's glass. Being polite, you fill your own glass last, but there's only enough left to fill it half-way. "Oh well," you say, "at least that's better than nothing." In this case, we might say, "nothing" refers, not strictly to

<sup>7</sup>Since  $v$  is additively separable, we have  $g(x, y) = |(f_1(x) + f_2(y)) - (f_1(0) + f_2(y))| = |f_1(x) - f_1(0)|$  and  $b(x, y) = |(f_1(x) + f_2(y)) - (f_1(x) + f_2(0))| = |f_2(y) - f_2(0)|$ . But because  $v$  is pleasure-friendly and pain-averse, we have  $|f_1(x) - f_1(0)| = f_1(x) - f_1(0)$  and  $|f_2(y) - f_2(0)| = f_2(0) - f_2(y)$ . It follows that  $g(x, y) - b(x, y) = (f_1(x) - f_1(0)) - (f_2(0) - f_2(y)) = f_1(x) + f_2(y) - (f_1(0) + f_2(0)) = v(x, y) - k$ .

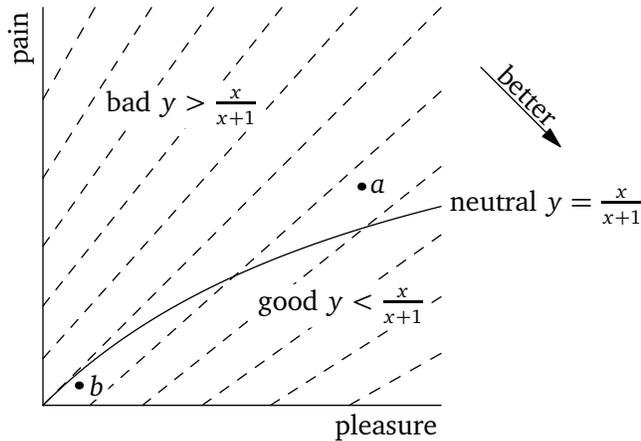


Figure 5: the second theory of well-being combined with the revised balanced analysis

*nothing*, but rather to a glass which contains nothing, to an empty glass, in other words.<sup>8</sup> A half-full glass is not as good as a full glass, but it’s still better than an empty glass. Similarly, we might think, when we say that a life is “better than nothing”, we mean that it is better than an empty life.

But how are we to understand “empty life”? In the context of a theory of well-being such as those discussed above, wherein the value of a life is thought to be a function of quantities of good and bad stuff contained in it, a natural answer is that an empty life is one which contains *zero* quantities of good and bad stuff [Cf. 2, p. 169]. In the model adopted above, we might call  $(0, 0)$  the empty life. Then we have the following analysis:

*The Empty Life Analysis.* A life is good iff it is better than an empty life.

If all empty lives are equally good, and the “better than relation” among lives is transitive, then this analysis implies that all good lives are better than all bad, so it is immune to objections of the sort made against the balance analysis above.<sup>9</sup>

A problem with it, however, is pointed out by John Broome [2, p. 170]. It implies that an empty life must be neutral. This might coherently be denied. One might say that a life is bad unless the good in it exceeds the bad by some positive margin, e.g., that  $(x, y)$  is good only if  $x - y > 1$ . On this view, a mere surplus of good over bad is not enough to make the life good; the surplus must be sufficiently large. Again, this seems a substantive moral issue. We should not prejudge it by our choice of analysis.

Notice, the Balance Analysis also has this problem. Both versions of this analysis entail that an empty life cannot be bad. If a life contains *no* bad, then it surely cannot contain *more* bad than good.

#### 1.4 The Worlds Analysis

In the analyses considered so far, lives have been conceived as containers, judged good or bad on the basis of what they contain. An alternative conception sees lives rather as

<sup>8</sup>Notice that in the expression “a glass which contains nothing”, “nothing” acts as a quantifier — though presumably one with restricted scope, since a glass which does contain something, e.g., some air, might still satisfy this expression in normal contexts.

<sup>9</sup>Strictly, we need another assumption here, viz., that there exists at least one empty life. I shall discuss this issue below, in relation to the Flatline Analysis (Section 2.2).

contents. Lives are contained in worlds. If a person exists in a given possible world, and she has a life there, then her life is, in a natural sense, contained in that world. She may have different lives in other worlds where she exists. Different worlds may contain different lives. Earlier I suggested that pleasure is good because it makes lives better by its presence. Likewise, we might say, a good life is good because it makes worlds better by its presence. A life is better than nothing, and hence good, just in case any world in which a person has this life is better than a world in which this person has no life, doesn't exist, other things being equal [2, p. 167].

This can be made more precise as follows. If  $w$  is a possible world and  $l$  a life, then say that another world  $w^*$  results from “merely adding”  $l$  to  $w$  iff (a) every person who exists in  $w$  also exists and has the same life in  $w^*$ , and (b) one additional person exists in  $w^*$  but not  $w$ , and the life this person has in  $w^*$  is  $l$ . Then we have the following analysis:

*The Worlds Analysis.* A life is good iff merely adding it to a world makes the world better.<sup>10</sup>

A problem with this is that it implies that certain well-known and apparently coherent theories in population ethics are actually incoherent. Critical Level Utilitarianism, for instance, avoids the Repugnant Conclusion by saying that lives which are “barely worth living” (and hence worth living) contribute negatively to the value of worlds which contain them [1, pp. 137–8]. Merely adding such a life to a world makes the world worse, not better. In order to make a positive contribution, a life must attain a certain “critical level”, located somewhere strictly above the point of being barely worth living.

Even if we don't accept this view, we can at least understand it. We comprehend the idea that a life which is good might nonetheless fail to contribute positively to the value of a world. Whether the mere addition of a good life always improves a world is a substantive issue. So it cannot be the case that a good life is by definition a life that makes a positive contribution.

What is needed here, you might think, is the familiar distinction between *personal* and *general* good [4, pp. 65–66]. Personal good, or value, is a matter of good things are for particular people. General good, or value, is a matter of how good things are overall, not for anyone in particular. When we say that a life is good, or worth living, we make a claim of personal good: we mean that the life is good for the one who lives it. However, such a life might not be good in the general sense. It might not make the world better by its presence. This is what proponents of Critical Level Utilitarianism think about lives which are barely worth living: these lives are good personally but not generally.

We may thus amend the analysis:

*The Worlds Analysis (Personal Version).* A life is good iff merely adding it to a world makes the world better *for the additional person*.

However, the problem with this, in my view, is that it gets things entirely backwards. The personal value of worlds is to be understood in terms of the value of lives, not the other way around. Contrast two questions, about the comparative value of worlds  $w$  and  $w^*$ :

1. Is  $w$  better than  $w^*$ ?

---

<sup>10</sup>More exactly: life  $l$  is good iff, for any worlds  $w$  and  $w^*$ , if  $w^*$  results from the mere addition of  $l$  to  $w$ , then  $w^*$  is better than  $w$ .

## 2. Is $w$ better than $w^*$ for Joanne?

The first is a question of general good, the second of personal good. Not all considerations relevant to the first question are relevant to the second, else the general-personal distinction would collapse. The considerations relevant to the question of personal good are only those which concern in some way how things are for Joanne, those which concern Joanne's *life*, in other words. But suppose now that Joanne does not exist, hence has no life, in  $w^*$ . Then there is no basis for comparing  $w^*$  with  $w$  in relation to Joanne's good. Neither world can be better or worse for Joanne.

I am arguing, in effect, for the view that existence cannot be either a benefit or a harm; a world where a person exists cannot be either better or worse for her than one where she doesn't exist. This view is controversial [see, e.g., 6, 9]. But I needn't insist on it here. My point, rather, is just this. Were the Worlds Analysis correct, then the view I have defended would entail that no lives are either good or bad. Intuitively, though, it does not entail this. Even those who reject this view, who think existence *can* be a benefit or harm, seem to acknowledge that whether a person's life is good is independent of whether she is benefited by her existence. Consider, e.g., this passage from Jens Johansson:

If your life is on the whole good, it may be natural to think that your existence is better for you than your non-existence. That is, of the two alternatives, coming into existence and never coming into existence, the former is better for you than the latter. [9, p. 285]

This suggests that even once we've settled that a person's life is good, the question remains open whether this person is benefited by her existence. Were the Worlds Analysis correct, however, then this question would be closed.

## 2 Temporal Analyses

### 2.1 Introducing Time

I turn now to temporal analyses. For simplicity, I shall continue to assume a hedonistic view of well-being. This is not strictly necessary; the temporal analyses I will consider are perfectly compatible with non-hedonistic views. But it will make things easier. The chief advantage is avoiding the need to give a general characterisation of what a life is.

To accommodate temporal analyses, a dimension of time will need to be introduced into our formal model. I do this as follows. Let a *life* now be a triple  $l = \langle d, \{x_t\}_{t \in (0,1]}, \{y_t\}_{t \in (0,1]} \rangle$ , where  $d$  is a positive real number representing the total duration of the life, in, say, years; and, for each  $t$  in the interval  $(0, 1]$ ,  $x_t$  and  $y_t$  are non-negative real numbers representing, respectively, the total pleasure and pain experienced during the first  $td$  years of the life. For example, if the life spans 100 years (i.e.,  $d = 100$ ), then  $x_{1/2}$  is the total pleasure experienced during the first 50 years,  $y_{3/4}$  is the total pain experienced during the first 75 years, and so on. Notice,  $x_0$  and  $y_0$  are undefined. This is because no pleasure or pain can be experienced, I assume, in a period of zero duration. Even the most fleeting of experiences must consume at least *some* time.

I assume, further, that all lives satisfy the following condition:  $\lim_{t \rightarrow 0} x_t = \lim_{t \rightarrow 0} y_t = 0$ . That is, if we consider ever shorter initial periods of a life (by "initial period" I mean simply a period beginning when the life begins), then, as the duration of these periods approaches zero, so too do the quantities of pleasure and pain they contain. This seems very plausible. One might argue for it as follows.

The amount of pleasure experienced in a given period is constrained by its length: the shorter it is, the less pleasure it can hold. This is not to say, of course, that longer periods always contain more pleasure than shorter ones, but only that longer periods have, as it were, a greater *capacity* for containing pleasure. Consider, for instance, a 100-year stretch fully saturated with pleasure, containing as much pleasure as could possibly be packed into 100 years. It seems inconceivable that this much pleasure could be condensed into a period of only, say, one minute (or one second, or one millisecond . . .). This quantity of pleasure exceeds the maximum that can be experienced in such a short space of time. Moreover, it seems plausible that the maximum quantity of pleasure that can possibly be experienced in a period will approach zero as the length of the period approaches zero. And this implies that  $\lim_{t \rightarrow 0} x_t = 0$ . (And a parallel argument may be given in the case of pain.)

Finally, I introduce the notion of “truncating” a life, by which I mean, roughly, removing a segment from the end of the life. More precisely, I define it as follows. Let  $l$  and  $l'$  be lives, with  $l' = \langle d', \{x'_t\}_{t \in (0,1]}, \{y'_t\}_{t \in (0,1]}\rangle$ . And let  $t$  be a number in  $(0, 1]$ . Then say that  $l$  is a *t-length truncation* of  $l'$  iff:  $t = d/d'$ , and for each  $i$  in  $(0, 1]$ ,  $x_i = x'_{ti}$  and  $y_i = y'_{ti}$ . It follows that, for each  $t$  in  $(0, 1]$ ,  $l$  has exactly one *t-length truncation*, which I shall call  $l_t$ . Thus, for example,  $l_{1/2}$  is the first half of  $l$ ,  $l_{1/3}$  the first third of  $l$ , and so on. Notice, there is no shortest truncation of  $l$ ; for every truncation, there is a shorter one. In particular, there is no 0-length truncation. A life of zero duration, I assume, would be no life at all. Every life — and hence every truncation of a life — must have positive duration.

I assume, as before, that the value of lives may be represented by a value function  $v$ , on a cardinal scale.

## 2.2 The Flatline Analysis

The first temporal analysis, which I call the “Flatline Analysis”, is largely similar to an analysis suggested by Broome [4, pp. 67–68, see also 1, pp. 24–25].

Here’s the basic idea. Consider a life all of whose truncations are equally good (i.e.,  $v(l) = v(l_t)$  for all  $t$  in  $(0, 1]$ ). Such a life has, in a certain sense, constant value over time. It never gets better or worse. To isolate precisely the sense I intend here, an analogy will be helpful.

Suppose you and I compete in a running race. We run differently. You maintain a steady pace throughout, whereas I start out faster, building up a lead over the first half of the course, then, tiring, slow down over the second half, allowing you to gradually catch up until, in the end, we finish in a dead heat. Now, focus on a time roughly three quarters through the race. Who is doing better in the race at this time?

Two answers seem possible. First: I am doing better, because I am in the lead. Second: you are doing better, because you are running faster. Neither answer is wrong. They simply use “better” in different senses. The first sense tracks the total distance accumulated by a runner at a given time. The second tracks the rate at which distance is being accumulated. In the first sense, my performance constantly gets better as the race goes on, because my accumulated distance gets greater (I get closer to the finishing line). In the second sense, my performance gets worse after half-way, because my rate of accumulation falls away (I slow down).

When I say that a life of the kind described above never gets better or worse, I mean this in the former, cumulative sense. The cumulative value of a life at a given time may be equated, I assume, with the value of the truncation of the life which ends at this time. The total value of a life up to time  $t$  is the total value it would have were it to end at  $t$ .

If all the truncations of a life are equally good, then the cumulative value of the life is, therefore, constant. For this reason, I call such a life a *flatline life*. Were we to represent its value over time on a graph, this would be represented by a flat, horizontal line.

A flatline life is “constantly neutral”, as Broome puts it [3, p. 67], neutral at every time. One might think that if a life is neutral at every time, then the life as a whole must likewise be neutral. And once we have identified one neutral life, we may then use it to define a good life by comparison with it, on the principle that a good life is better than a neutral one. This line of thought leads to the following analysis.

*The Flatline Analysis.* A life is good iff it is better than a flatline life.

On this approach, the “nothing” in “better than nothing” refers to a flatline life. This seems a natural enough interpretation. Recall the running analogy: a runner whose cumulative distance is the same at every time during the race is one who never steps off the starting line. Likewise, we might say, a flatline life is one that “goes nowhere”. In a flatline life, the rate at which value accumulates is constantly zero, nothing.

### 2.3 The Limit Analysis

The second temporal analysis, which I call the “Limit Analysis”, is essentially the same as an analysis suggested by Charles Blackorby, Walter Bossert and David Donaldson [1, p. 25].

To see the rationale for this analysis, it will be instructive to think again about the expression “better than nothing”. Consider a typical use of this phrase. You really wanted to have gravy on your chips but there was only ketchup. “Oh well,” you say, “ketchup isn’t the best, but at least it’s better than nothing.” But here the comparison is really between one thing and another, not between something and nothing. There is an implicit *baseline*, and the thing said to be “better than nothing” is a certain alteration or addition to that baseline. In the example, the baseline is the chips alone, and what you mean by saying that ketchup is “better than nothing” is that adding to the baseline, by putting ketchup on the chips, is better than leaving it as it is. Or, in other words, chips with ketchup is better than chips on their own. Roughly, then, “ $x$  is better than nothing” means that baseline-plus- $x$  is better than baseline-plus-nothing, i.e., the baseline alone.

An implicit baseline analysis of this sort is well suited to explaining what it is for smaller parts of lives, in particular, *futures* of lives, to be worth living. Suppose a woman loses her legs in a car accident. This is a tragedy, to be sure; nonetheless, we might think, she was lucky to have escaped death. Though her future holds great adversity, it’s still better than nothing. Here the implicit baseline is that part of the woman’s life which she has lived already, up to the time of the accident. We judge that adding to this baseline a future of adversity is better than leaving it as it is, with no future attached. The whole life, including the future, is judged better than the life that would remain were the future removed [Cf. 1, p. 23].

To forestall objections, let me make two clarifications. First, on the proposed analysis, a future is worth living just when it has positive *marginal* value, i.e. when its addition produces an increment in value. But this is not to say that the future has positive *intrinsic* value, that it should be judged good when considered in isolation, as a whole life in itself. Consider a person whose life is generally happy, except for the final five years, which is misery, because he contracts a disease which slowly and painfully kills him. The only

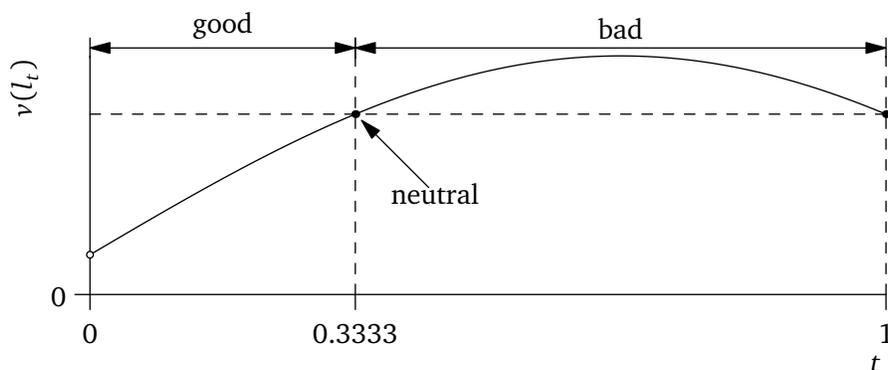


Figure 6: when  $l$  has a good, bad, or neutral future

redeeming part of this last period is his finally fulfilling some major lifelong goal: he proves an important mathematical theorem he has worked on for many years. Suppose we judge that removing the final five years would overall have made his life worse, for although this would have saved him considerable pain and misery, it would also deprive him of achieving one of his greatest ambitions, and the former gain would not be sufficient to outweigh the latter loss. On the proposed analysis, then, he had a future worth living at the time when he first got sick, before proving the theorem.

Now imagine another person whose *entire* life is intrinsically identical to the last five years of the first person's life. We might nonetheless judge that this person's life is not worth living, for in this case, the proving of the theorem is no longer such a big deal, because it's not the attainment of a lifelong goal, the culmination of years of hard work. The latter judgement might seem to contradict the former one, i.e. that the first person had a future worth living, but this relies on a picture of value we needn't accept, a sort of atomism according to which the value of a whole must be the sum of the values of its parts considered in isolation.<sup>11</sup> At any rate, such atomism is no part of the analysis I am considering.

Second, to say that your life does not have a future worth living is not to say that you ought to end it now, or that it would be irrational for you to continue living if it is possible for you to avoid this. This doesn't follow even on the assumption that what is rational for you to do is whatever will result in your having the best life you can have.<sup>12</sup> It might be that though the future is not worth living, part of it is, and you might be able to live part without living all.

In terms of the formal model introduced above, we have thus have the following:

*The Limit Analysis (Futures).* The future of  $l$  at  $t$  is good iff  $v(l) > v(l_t)$ .

This is illustrated in Figure 6. The horizontal dashed line represents the value of the whole life. Thus points on the curve above this line correspond to times when the future is good, points below to times when it's bad, and the solitary point where the two intersect (at  $t = 1/3$ ) to a time when it's neutral.

Not much more is required now to extend this analysis to the case of whole lives. Consider a time very early in the life, say  $t = 0.01$ . At this time, nearly all of the life lies

<sup>11</sup>Cf. 11, on the general issue of atomism with respect to value see 5.

<sup>12</sup>This assumption is controversial. See e.g. [10, Part 2].

in the future. The future of the life is almost the same thing as the whole life, and so the future's being good is almost the same as the whole life's being good. But still, it's not quite the same. One might say that the future's being good at this time *approximates* the whole life's being good. We could take an even earlier time, say  $t = 0.001$ , and get a more accurate approximation. Still, the possibility would remain that the future is good while the life is not (or *vice versa*). This possibility might become increasingly remote as  $t$  gets smaller, but never vanishes entirely.

What we need to do, you might think, is consider a time so early that the corresponding baseline would be an “infinitesimal” truncation, a truncation shorter than all the rest, yet whose duration is non-zero. This is not in fact possible, in the model I've adopted, as there is no such truncation. Nonetheless, using the familiar tools of calculus, we can understand statements about infinitesimal truncations without assuming the existence of any. The key, of course, is the mathematical concept of a limit. What we mean by the value of an infinitesimal truncation of a life, we may say, is the *limit* of the value of a truncation of this life as its duration *approaches* zero, i.e.,  $\lim_{t \rightarrow 0} v(l_t)$ . In Figure 6, this limit is represented by the hollow dot at  $t = 0$ . Since this is below the horizontal dashed line, we may say, the life as a whole is good.<sup>13</sup>

Thus we have:

*The Limit Analysis.*  $l$  is good iff  $v(l) > \lim_{t \rightarrow 0} v(l_t)$ .

One could perhaps gloss this by saying that a life is good just in case it is better than an infinitesimal truncation. And this might seem to provide an answer to our earlier question: What does “nothing” refer to in the predicate “. . . is better than nothing” when applied to a life? The Limit Analysis might seem to entail that it refers to an infinitesimal truncation. But this is not really so. The gloss is fine, but only if one remembers it is mere shorthand for the more complicated description which makes no reference to infinitesimal things of any sort. It cannot be right that “nothing” refers to an infinitesimal truncation, because there is no such thing. On this analysis, we may say that “nothing” refers to nothing. That is, in this context, “nothing” is not a referring term. The predicate “. . . is better than nothing”, when applied to whole lives, behaves as a sort of semantic organic unity, whose meaning is not composed simply out of the meanings of its parts. Asking what “nothing” refers to in, say, “His life was better than nothing” is in this respect like asking what “bucket” refers to in “He kicked the bucket”.

What if the relevant limit is undefined, doesn't exist? In that case, I assume, sentences like “ $v(l) > \lim_{t \rightarrow 0} v(l_t)$ ” would be false, and so the Limit Analysis would imply that such a life is neither good, bad, nor neutral. We might think of this as a case of incommensurability: the life is neither better than, nor worse than, nor equal in value with nothing. This is at least consistent. Saying that a life is neither good, bad, nor neutral is not like saying, e.g., that a life is both good and bad. The latter is, I think, genuinely inconsistent.<sup>14</sup>

<sup>13</sup>I have heard the objection that the Limit Analysis is too complicated or mathematically sophisticated to give the meaning of our ordinary concept of a good life. The correct analysis of an ordinary concept, so the objection goes, cannot be one that employs a non-ordinary concept like that of a mathematical limit. However, I don't think this can be right. For example, instantaneous speed is an ordinary concept; normal folks are not baffled by the speedometers in their cars. But speed cannot be properly analysed without the concept of a limit (or something at least as mathematically sophisticated, like hyperreal numbers, as in non-standard analysis).

<sup>14</sup>I take the following to be conceptual truths: (a) all good lives are better than all bad; and (b) nothing is better than itself. Thus, the proposition that  $l$  is both good and bad entails that  $l$  is both good and not good.

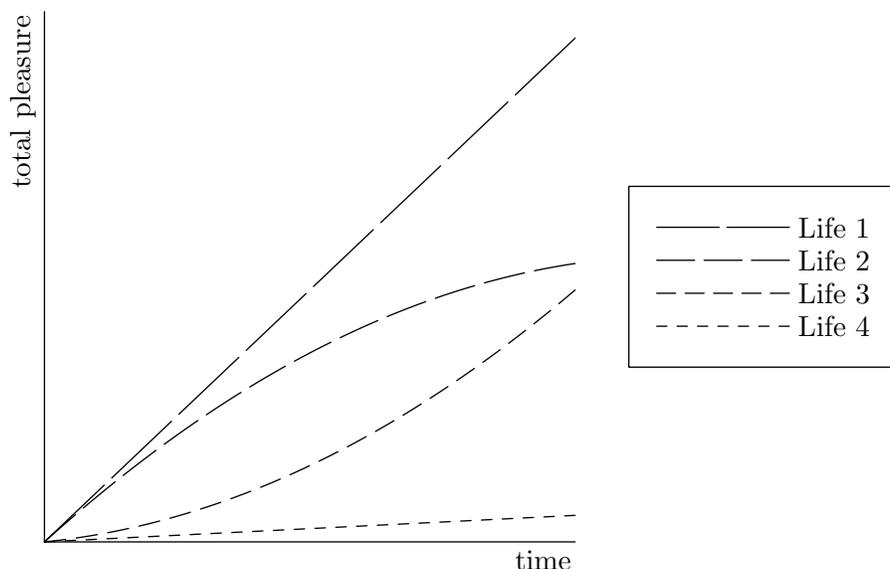


Figure 7: total pleasure in four lives

## 2.4 Advantages of the Temporal Analyses

The temporal analyses do not share the problems of the non-temporal analyses discussed above. The problem with the Balance Analysis, as we saw, is that when combined with certain value functions it has the unwelcome consequence that some good lives are worse than some bad lives. The two troublesome value functions considered above present no such difficulties for either the Flatline Analysis or the Limit Analysis.

The problematic value function for the first version of the Balance Analysis was  $v(x, y) = x - 2y$ , or, in terms of our new temporal framework,  $v(l) = x_1 - 2y_1$ . Given this function, it follows from the definition of  $l_t$  that  $v(l_t) = x_t - 2y_t$ . So  $\lim_{t \rightarrow 0} v(l_t) = 0$ , because, recall,  $\lim_{t \rightarrow 0} x_t = \lim_{t \rightarrow 0} y_t = 0$ . And this implies that  $v(l) = 0$  for any flatline life  $l$ , because if  $v(l_t) = v(l)$  for all  $t \in (0, 1]$ , then  $\lim_{t \rightarrow 0} v(l_t) = v(l)$ . Both temporal analyses imply, therefore, that  $l$  is good only if  $v(l) > 0$ , and bad only if  $v(l) < 0$ ; so a good life must be better than a bad life.

The problematic value function for the second version of the Balance Analysis was, in our new framework,

$$v(l) = \frac{x_1 + 1}{y_1 + 1}.$$

Similar reasoning shows that, combined with either temporal analysis, this implies that  $l$  is good only if  $v(l) > 1$ , and bad only if  $v(l) < 1$ . So, again, good is better than bad, as it should be.

Other value functions, however, are not so friendly to the temporal analyses. Here is one example:  $v(l) = (x_1 - y_1)/d$ . This measures the value of a life by its average balance of pleasure over pain per year of duration. To see the trouble this causes for the temporal analyses, consider the four lives represented in Figure 7. For simplicity, assume that all lives have the same duration and that none contains any pain. Then the value of these lives, according to this value function, is shown in Figure 8. Life 2 is worse than Life 1, which is a flatline life. And an “infinitesimal truncation” of Life 2 is better than the whole life. Thus Life 2 is bad according to both the Flatline Analysis and the Limit Analysis

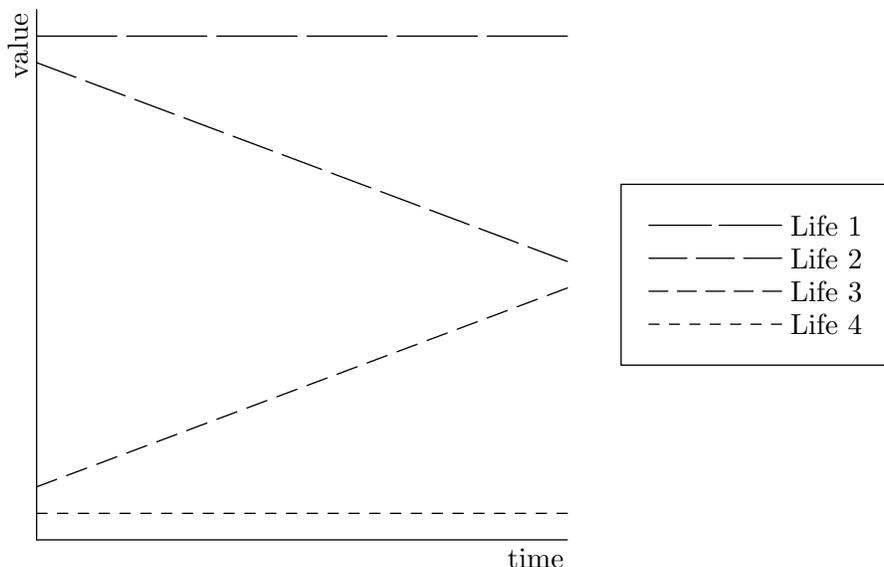


Figure 8: value of the four lives, assuming  $v(l) = (x_1 - y_1)/d$

(given this value function). Similarly, Life 3 is good according to both analyses. But Life 2 is better than Life 3. So again we have the absurd result that a bad life is better than a good life.

In this case, however, I think the best response might be to keep the analysis, and to reject the incompatible value function instead. This value function is, I shall argue, independently quite implausible.

A widely held view is that value supervenes on other properties. There could not be a *bare* difference in value, i.e., two things that differ in value alone, not in any other way. This applies to lives as much as to anything else. If two lives are the same in all other respects, then they must be equally good. A natural extension of this is that the extent to which lives are able to differ in value — the extent to which one life can be better or worse than another — is constrained by the degree of dissimilarity between the lives in other respects. Lives which are otherwise only very slightly dissimilar cannot be vastly dissimilar in value. But it seems plausible that the maximum possible degree of dissimilarity between lives of a given duration will approach zero as their duration approaches zero. Lives of only, say, one minute cannot differ more than very slightly; and lives of only one second cannot differ even that much. Were we to consider a sequence of lives ever decreasing in duration, then as the lives became vanishingly short, the differences between them would likewise become vanishingly slight. It follows that the difference in *value* between lives of a given duration must equally approach zero as their duration does so. That is, a plausible value function, according to this argument, must satisfy the following condition: for any lives  $l$  and  $l'$ ,  $\lim_{t \rightarrow 0} [v(l_t) - v(l'_t)] = 0$ .

But the value function above — the problematic one for the temporal analyses — does not. Focus on Lives 1 and 4 from the previous example. As shown in Figure 7, if we trace these lives backwards in time, starting at the end and working back to the beginning (i.e., moving from right to left on the graph), we see the difference between the lives steadily shrinking. According to this value function, however, there is no such convergence in the *value* of these lives. As shown in Figure 8, the difference in value is constant.

If a value function satisfies the above convergence condition, then it is compatible with both temporal analyses. Suppose, first, that  $l$  is good and  $l'$  bad according to the Flatline Analysis. Then there exist flatline lives  $f$  and  $f'$  such that  $v(l) > v(f)$  and  $v(l') < v(f')$ . But, given convergence, all flatline lives are equally good, and so  $v(f) = v(f')$ . Hence,  $v(l) > v(l')$ . Now suppose  $l$  is good and  $l'$  bad according to the Limit Analysis. Then  $v(l) > \lim_{t \rightarrow 0} v(l_t)$ , and  $v(l') < \lim_{t \rightarrow 0} v(l'_t)$ , but given convergence,  $\lim_{t \rightarrow 0} v(l_t) = \lim_{t \rightarrow 0} v(l'_t)$ . Hence  $v(l) > v(l')$ .

The ideal, as I said earlier, would be to find an analysis compatible with *all* value functions. But, failing achievement of this ideal, the cost of excluding some class of value functions is, I suggest, diminished the more implausible the class so excluded. And, as we have just seen, non-convergent value functions are, in certain way, quite implausible. They seem to allow the value of lives to float free from their other properties. So, excluding them seems not so great a cost for the temporal analyses. Earlier, we saw that the Balance Analysis would require us to exclude a different class of value functions: those which are not additively separable. This seems comparatively more costly. I know of no similar rationale for requiring additive separability. And in any case, as was noted above, the Balance Analysis is unsuitable for those theories of well-being which, like preference-based theories, are not defined in terms of good and bad stuff. So the temporal analyses seem preferable to the Balance Analysis.

They also seem preferable to other two non-temporal analyses. The problem with the Empty Life Analysis, as we saw, is that it implies that empty lives must be neutral, whereas it seems possible to hold that such lives are bad. The temporal analyses do not share this problem. Consider, for example, the value function  $v(l) = d(x_1 - y_1 - 1)$ . Combined with either temporal analysis, this implies that  $l$  is good iff  $x_1 - y_1 > 1$ , and so if  $l$  is empty (i.e.,  $x_1 = y_1 = 0$ ), then  $l$  is bad.<sup>15</sup> The problem with the Worlds Analysis is that it conflates the question whether a person's life is good or bad with the question whether she is benefited by her existence. This is clearly not a problem for either temporal analysis, because they do not involve comparisons between worlds at all.

## 2.5 Flatline vs Limit

Which of the two temporal analyses is preferable? Here there is not much to decide either way. Each analysis has a slight disadvantage compared to the other.

The disadvantage of the Limit Analysis is that it requires that the goodness of lives may be represented numerically, whereas the Flatline Analysis requires only a ranking of lives. Some theories of well-being supply only a ranking, not a numerical function. Consider, for example, the hedonist view according to which avoiding pain has lexical priority over achieving pleasure. This provides an ordering of lives:  $l$  is better than  $l'$  iff either (a)  $y_1 < y'_1$  or (b)  $y_1 = y'_1$  and  $x_1 > x'_1$ .<sup>16</sup> But this ordering cannot be represented by a numerical function, and so this theory of well-being is incompatible with the Limit Analysis. On the other hand, the definition of a flatline life requires only an ordering. Given the above ordering, a flatline life is one all of whose truncations contain zero pleasure and zero pain. So, combined with the Flatline Analysis, this theory of well-being implies that a life is bad iff it contains any pain at all. But, of course, this theory of well-being is very implausible

<sup>15</sup>We have  $v(l_t) = dt(x_t - y_t - 1)$ , and so  $\lim_{t \rightarrow 0} v(l_t) = 0$ . So  $l$  is good, according to either temporal analysis, iff  $d(x_1 - y_1 - 1) > 0$  iff  $x_1 - y_1 > 1$ .

<sup>16</sup> $l$  and  $l'$  are equally good iff  $l$  is not better than  $l'$  nor  $l'$  better than  $l$ .

(surely a mild pain can be compensated for by a sufficiently great pleasure), as, I would say, are all lexical priority views. Excluding them seems, therefore, not a very great cost.

The disadvantage of the Flatline Analysis is that it is incompatible with the possibility that every possible life is good. To see this, suppose that every life is good. Then no life is neutral, and so no life is flatline. (Surely it is no promising strategy to reject this step in the argument by holding that a life may be both good and neutral.) But if there is no flatline life, then no life is better than a flatline life, and so no life is good. Contradiction. The view that every possible life is good is not very plausible; indeed, it may appear morally perverse given the depths of misery and suffering that some people experience in their lives. So excluding this view is perhaps not a great cost. But it is still a cost, I think, if only a small one. And it is a cost not shared by the Limit Analysis. To take a simple example, consider the value function  $v(l) = d$ . This says that the value of a life depends solely on its length: the longer the better. Combined with the Limit Analysis, this implies that every life is good.

As we have seen, however, the two temporal analyses tend to agree in a broad range of cases. Perhaps, then, deciding between them is not so urgent.

## References

- [1] Charles Blackorby, Walter Bossert and David Donaldson. *Population Issues in Social Choice Theory, Welfare Economics, and Ethics*. Cambridge University Press, 2005.
- [2] John Broome. ‘Goodness is Reducible to Betterness: The Evil of Death is the Value of Life’. In: *Ethics Out of Economics*. Cambridge University Press, 1999.
- [3] John Broome. *Weighing Goods*. Blackwell, 1991.
- [4] John Broome. *Weighing Lives*. Oxford University Press, 2004.
- [5] Campbell Brown. ‘Two Kinds of Holism about Values’. In: *The Philosophical Quarterly* 57.228 (2007), pp. 456–463.
- [6] Krister Bykvist. ‘The Benefits of Coming into Existence’. In: *Philosophical Studies* 135.3 (2007), pp. 335–362.
- [7] Lewis Carroll. *Alice’s Adventures in Wonderland, and Through the Looking Glass*. Penguin Classics, 1998.
- [8] Thomas Hurka. ‘Asymmetries in Value’. In: *Noûs* 44.2 (2010), pp. 199–223.
- [9] Jens Johansson. ‘Being and Betterness’. In: *Utilitas* 22.3 (2010), pp. 285–302.
- [10] Derek Parfit. *Reasons and Persons*. Oxford University Press, 1984.
- [11] J. David Velleman. ‘Well-Being and Time’. In: *The Metaphysics of Death*. Ed. by John Martin Fischer. Stanford University Press, 1993.