The Value of Life Itself

_____

Andrew Y. Lee
Rice University, Dept. Of Philosophy
http://andrewyuanlee.com/
andrewyuanlee@gmail.com

ABSTRACT:

There are some who think that life is worth living not merely because of the goods and the bads within it, but also because life itself is good. This paper argues instead that life itself is neither good nor bad. I explain how the view that life itself is good yields a dilemma: either (1) any finitely long excellent life is worse than some longer life wholly devoid of pleasure, desire-satisfaction, knowledge, or any other goods, or (2) very short lives filled with nothing but pain are worth living. Since neither result is acceptable, we ought to reject the view that life itself is good. The resulting picture has implications for questions about the threshold for a life worth living, the value of consciousness, the basic welfare goods, a variety of applied ethical issues concerning life and death, and whether death is bad.

INTRODUCTION

Consider the idea that the goodness of a life is determined not only by the goods within that life, but also by the fact that life itself is good:

> There are elements which, if added to one's experience, make life better; there are other elements which if added to one's experience, make life worse. But what remains when these are set aside is not merely neutral: it is emphatically positive. Therefore life is worth living even when the bad elements of experience are plentiful, and the good ones too meager to outweigh the bad ones on their own. The additional positive weight is supplied by experience itself.
> —Thomas Nagel, Mortal Questions [1979]

On this view, even after measuring all the goodness and badness from within a life, there remains some extra goodness from life itself. Because of this, a life where the severity of the bads outweigh the benefit of the goods may nevertheless be worth living. Imagine a scale that weighs the good against the bad, but where the plate weighing the good is itself heavier: to achieve equilibrium, the weight from the bads must exceed the weight from the goods. The result is that the goodness of a life is always greater than the goodness solely due to the goods from within that life, for life itself is good.

This paper argues, by contrast, that life itself is neither good nor bad. My core argument is that the view that life is worth living for its own sake yields a dilemma: either (1) any finitely long excellent life is worse than some longer life wholly devoid of pleasure, desire-satisfaction, knowledge, or any other goods, or (2) very brief lives full of nothing but horrible pain are worth living. Since neither result is acceptable, we ought to reject the view that life is worth living for its own sake. On the view I favor, a life may be worth living because of the goods that it contains, but never because life itself is a good. In brief: life is neutral.

The question of whether life is worth living for its own sake impacts a variety of other philosophical issues. These include issues about the threshold for a life worth living, the basic welfare goods, the grounds of moral status, the value of consciousness, and the ethics of euthanasia, abortion, suicide, procreation, murder, and death. These connections will be briefly discussed at the end of the paper.

This paper has five sections. In **§1**, I develop some conceptual foundations for investigating the value of life. In **§2** and **§3**, I present and develop my core argument, which I call the 'Argument for Zero'. In **§4**, I respond to various ways of resisting the Argument for Zero. In **§5**, I briefly discuss implications for other philosophical issues.

§ 1 | LIFE

Let me begin with some basic stipulations about the notion of a *life*. For any life, we can ask how good or bad that life is, whether that life is worth living, or whether that life is better or worse than some other life. The *character* of a life (which we may simply take to be the set of descriptive facts about that life) determines the *value* of that life (or how good or bad that life is). Better lives yield more *positive* values while worse lives yield more *negative* values. A life is *worth living* just in case its value is greater than *zero*.

Let us say that the *global value* of a life is how good or bad that life is overall. For most of this paper, I will think of the global value of a life as the sum of two quantities: (1) the value due to the goods and bads within the life, and (2) the value due to life itself. I will later consider views that take other factors to contribute to the global value of a life, and I will later consider the possibility that what comparisons between lives depend on factors besides global value. But taking these points for granted for the moment will simplify the discussion.

The value from *within a life* is determined by the set of goods (and bads) that the life contains. By a *good*, I mean something that makes a life intrinsically better (and by a *bad*, I mean something that makes a life intrinsically worse). Which kinds of things are good will depend on one's preferred theory of welfare, but common candidates for goods include pleasure, desire-satisfaction, and knowledge. My focus will be on the value due to the entire set of goods within a life, rather than the value due to any particular good. Because of this, I will remain neutral on whether or not the values of goods combine additively (meaning that the value generated by a set of goods is the sum of the value generated by each of those goods individually).

The value due to the goods and bads *within* a life may be distinguished from the value from *life itself*. The principal question of this paper is whether there is any value due to life itself. The *positive theory*, which is the view I sketched at the

beginning, holds that the value of life is positive.[1] The *neutral theory*, which is the view that I endorse, holds that the value of life is zero.[2] Everyone agrees that the goods and bads within a life generate value for that life: the issue is whether there is additional value from life itself. There is, of course, the conceptual possibility of a *negative theory*, which takes the value of life to be negative, but I will assume that such a theory is a non-starter. For brevity I will use the phrase 'value of life', though in contexts that are liable to ambiguity I will use the phrase 'value of life itself'.

These questions about the values of lives are fundamentally questions about welfare. What it is to have a life is to be a welfare subject, or to be the kind of thing that can be doing well or badly. The question of which entities have lives is the question of what makes an entity a welfare subject. The global value of a life is the total welfare generated by that life. The goods or bads within a life are the welfare goods contained within that life. A life is worth living just in case its welfare is above zero. And if life is worth living for its own sake, then life itself is a welfare good.[3]

When I talk about value, I mean value that is *intrinsic* (in that it is valuable for its own sake), *pro-tanto* (in that it is defeasible), and *personal* (in that it is for an individual). I take the term 'value' to be neutral between positive value and negative value, 'goodness' to be synonymous with 'positive value', and 'badness' to be synonymous with 'negative value'. I will remain neutral on questions about the nature of value (meaning questions at the level of metaethics). But I will take for granted that the values of lives can be represented by real numbers, where zero marks the threshold for a life worth living. As a result, I will assume that for any two lives A and B, either A is better than B or B is better than A or they are equally good. It is possible to generate a version of my argument without these assumptions, though taking them for granted simplifies the exposition.

THE BAD-LIFE-WORTH-LIVING INTUITION

Consider the following life:

---

[1] See Frankena [1973], Nagel [1979], Schumacher [2010, p. 204], and Kriegel [2019] for endorsements of the positive theory.

[2] See Glannon [2016] and Lee [2018] for endorsements of the neutral theory (at least about consciousness).

[3] See Campbell [2016] for a recent overview on theories of well-being.

—————————

BAD LIFE: A life whose average quality is negative.

—————————

By 'average quality', I mean a quantity that depends only on the goods and bads within life (and not on life itself). More specifically, the *average quality* of a life is the value due to the goods and bads within the life divided by the length of that life. If life itself is good, then any given life's average quality may come apart from its average value, since the goodness of life itself may add value beyond that which is generated by the goods and bads within the life. By contrast, if life itself is neutral, then average quality and average value are equivalent.

The *bad-life-worth-living intuition* is the intuition that some version of BAD LIFE is worth living.[4] Everyone can agree that if we consider only the goods and bads within BAD LIFE, then it would not be worth living. Yet some nevertheless have the intuition that BAD LIFE is worth living. The explanation offered by those who endorse this intuition is that there is some goodness from life itself that offsets the badness of the bads within BAD LIFE. Putting it another way, a proponent of the bad-life-worth-living intuition thinks that even though the average quality of BAD LIFE is negative, the average value of BAD LIFE is positive.

The bad-life-worth-living intuition is the core motivation for the positive theory. Some, such as Kriegel [2019], suggest only that a life of moderately bad quality would nevertheless be worth living. Others, such as Nagel [1979], seem to think that even a life of very bad quality would nevertheless be worth living. And some, such as Schumacher [2010], suggest any life whatsoever, no matter how bad, would be worth living. For the purposes of my arguments, it will not matter how exactly the positive theorist specifies the details of BAD LIFE. But let us stipulate that however

---

[4] See Nagel [1979], Schumacher [2010, p. 205], and Kriegel [2019] for endorsements of the bad-life-worth-living intuition. See Frankena [1973], Lamb [1998, p. 45], Agar [2001], Link [2013] for endorsements of the claim that life is intrinsically valuable. Note that some philosophers who claim that life is intrinsically valuable are focused mainly on impersonal value (whereas my focus is on personal value).

the positive theorist calibrates their intuition, the average value of BAD LIFE is just barely positive by their lights. In other words, the badness from the bads in BAD LIFE is just barely outweighed by the goodness from the goods (including life itself).

§ 2 | THE ARGUMENT FOR ZERO

The Argument for Zero appeals to the following kinds of lives:

An *excellent* life: a life with an average quality very far above zero.
An *awful* life: a life with an average quality very far below zero.
An *empty* life: a life devoid of any goods or bads (except for life itself).

To imagine an excellent life, think of the best points from your own life and imagine a life comprised filled with points that are much better. To imagine an awful life, think of the worst points from your own life and imagine a life filled with points that are much worse. To imagine an empty life, think of a life comprised of nothing but neutral experiences of gray, with no pleasures (nor pains), no desires (satisfied or frustrated), no knowledge (or even beliefs), and nothing else from a standard list of welfare goods (or bads).

Now consider the following lives:

_____

PARADISE: An excellent life that lasts 73 years.[5]

ETERNITY: An empty life that lasts indefinitely.

SWIFT HELL: An awful life that lasts one minute.

_____

---

[5] Why 73 years? For now, simply note that this is the present average length of a human life.

With these lives defined, I can present my core argument:[6]

The Argument for Zero[7]

P1:   PARADISE is better than ETERNITY.

P2:   SWIFT HELL is not worth living.

P3:   If life is worth living for its own sake, then either P1 or P2 is false.

C: Life is not worth living for its own sake.

The argument is valid. I think both P1 and P2 are extremely plausible, and I will simply take them for granted. The remaining premise is P3, which will require several stages of argumentation to adequately develop and defend.

Before moving forward, let me make a preemptive remark. Over the course of the paper, I will show how different classes of positive theories can be formalized via equations that relate the value of life to the length of life. These formalizations may initially strike some as unnecessary for evaluating the philosophical issues. But I wish to forestall that sentiment with a promise that they will yield philosophical dividends. As we will see, the formalizations will play a key role in illustrating why every version of the positive theory yields counterintuitive results.

THE ARGUMENT FROM ETERNITY

Any positive theorist must answer the following question: how does the value of life itself relate to the length of life? A natural answer is that more life means more value: the longer a life, the greater the value generated by life itself. This leads to the simplest and most straightforward version of the positive theory:

Linear: The value of life increases linearly as a function of the length of life.[8]

---

[6] Though my arguments appeal to lives of infinite length and infinite value, it is possible to reframe my arguments so as to appeal only to lives of finite length and finite value. I appeal to the infinitary cases because doing so simplifies some of the exposition.

[7] I will later develop a more sophisticated version of the Argument for Zero that is more general (though which remains simple in structure).

[8] Though I focus on *Linear*, my argument here will apply to any positive theory that takes the value of life to increase unboundedly as a function of the length of life.

If *Linear* is true, then it is straightforward to specify an equation that relates the value of life to the length of life. To do so, we need to introduce two functions, which I will call 'λ' and 'τ':[9]

λ(L) = the value of life itself for L
τ(L) = the length of life L (in years)

λ takes as input a life and outputs the value generated by life itself. τ takes as input a life and outputs the length of that life (in years). As we will see, these functions will reoccur many times over the rest of the paper. With λ and τ, we can capture *Linear* with a simple equation:

$$\lambda(\text{L}) = a \times \tau(\text{L})$$

The Linear Equation

The constant *a* allows us to differentiate theories that accept *Linear* but vary on how much value is generated by life itself for a life of a given length. In other words, *a* allows one to scale the value of life itself, where higher values of *a* correspond to positive theories that take life itself to generate greater amounts of value (given a fixed length of life).

Though *Linear* may seem attractive, it leads to the Argument from Eternity:

The Argument from Eternity[10]

P1:   ETERNITY has infinite value.
P2:   PARADISE has finite value.
C: ETERNITY is better than PARADISE.

---

[9] In the paper, I generally use Greek letters (such as λ) to denote functions, lowercase Latin letters (such as *a*) to denote constants, and small-caps Latin letters (such as L) to denote lives.
[10] Every argument I take to be unsound has '#'' in front of its title and is called the 'Argument *from* X' (rather than the 'Argument *for* Y' or the 'Argument *against* Z').

Here is the basic reasoning. ETERNITY lasts indefinitely. By *Linear*, the value of life increases linearly with the length of life. By the positive theory (and the bad-life-worth-living intuition), the value of life is positive. Since any positive real number times infinity is infinity, ETERNITY is guaranteed to generate infinite value from life itself. Since ETERNITY contains no bads, there is nothing to offset the goodness from life itself. Therefore, the global value of ETERNITY is infinite. By contrast, given that GOOD LIFE is only finitely long, it is plausible that the global value of GOOD LIFE is finite. This mean that if we accept P1 and P2, then ETERNITY is better than GOOD LIFE. But of course, that is the wrong result: it is obvious that GOOD LIFE is better than ETERNITY. Therefore, we ought to reject one of the premises in the Argument from Eternity. In particular, we should reject P1.

The Argument from Eternity has some similarities to the Repugnant Conclusion, the thesis that for any world A containing a finite number of excellent lives, there is another world Z containing a greater number of lives barely worth living such that Z is better than A. However, it would be a mistake to think that the Argument from Eternity is simply a repackaged version of the Repugnant Conclusion where the population variable is replaced with a time variable. Consider the fact that one *cannot* reject the stipulation in the Repugnant Conclusion that each person in Z has a life barely worth living (since that is simply how the scenario is defined). By contrast, one *can* reject the supposition in the Argument from Eternity that the average quality of ETERNITY is positive (since that is precisely the premise that the neutral theorist rejects). The Repugnant Conclusion generates a feeling of paradox because it is unobvious which premise leading to the conclusion ought to be rejected. By contrast, the Argument from Eternity generates no feeling of paradox because it is obvious where the argument goes wrong: namely, with the premise that ETERNITY has infinite value.[11]

---

[11] There are also other differences between the Argument from Eternity and the Repugnant Conclusion. For example, some responses to the Repugnant Conclusion, such as revising the notion of a life worth living, rejecting the transitivity of better-than, or appealing to person-affecting principles do not have any obvious analogues as responses to the Argument from Eternity. For more on the Repugnant Conclusion, see Arrhenius, Ryberg, & Tännsjö [2017].

Where does that leave us? Though *Linear* may have seemed attractive, the Argument from Eternity reveals that it yields unacceptable results. Let us move on to other versions of the positive theory that avoid this problem.

THE ARGUMENT FROM HELL

We began with the following question: how does the value of life itself relate to the length of life? To escape the Argument from Eternity, the positive theorist must reject *Linear* (and more generally, any theory that takes the value of life itself to increase unboundedly). A natural alternative is the following view:

Constant: The value of life is a constant.

According to *Constant*, the value of life itself is all or nothing: every life generates the same amount of value from life itself, regardless of its length. This view is captured by the following equation (where c is the constant value of life itself):[12]

$$\lambda(\text{L}) = c$$

The Constant Equation

By accepting *Constant*, the positive theorist can escape the Argument from Eternity, since it will no longer be the case that ETERNITY has infinite value. But we immediately run into a new issue: the positive theorist must qualify the bad-life-worth-living intuition. Since there is no limit on how long BAD LIFE might last, there is no limit on how much badness BAD LIFE might generate.[13] This means that

---

[12] I will assume that c is finite. Otherwise, the Argument from Eternity would rearise.

[13] At least, if we grant the plausible assumption that adding bads to a life can increase the badness of that life unboundedly. Suppose then that the average quality of BAD LIFE is just barely negative. Let φ be a function from a life L to the value due to the goods and bads within L. Let c be the maximal value generated by life itself. Since BAD LIFE can be arbitrarily long (and thus become arbitrarily bad), it is guaranteed that for some version of BAD LIFE, $\varphi(\text{BAD LIFE}) + c < 0$. This means that even if we allow the average quality of BAD LIFE to be arbitrarily close to zero, it is guaranteed that the bad-life-worth-living intuition fails to hold for lives of arbitrary length (so long as one takes the value of life itself has a maximum).

rejecting *Linear* requires taking the bad-life-worth-living intuition to have a limited scope, where not all bad lives are worth living, even if we hold fixed their average quality. That strikes me as a concession, since it is natural to interpret the bad-life-worth-living intuition as a restricted universal claim, scoping over all lives whose average qualities are equal to or greater than that of BAD LIFE.

The natural response for the positive theorist is to contend that the bad-life-worth-living intuition is intended to hold only at the scale of human lives, rather than for lives of arbitrary length. In fact, those who have endorsed the bad-life-worth-living intuition were probably thinking only about lives of the kinds that humans have. By restricting the scope of the bad-life-worth-living intuition, the positive theorist can retain the claim that PARADISE is better than ETERNITY while holding on to the idea that some versions of BAD LIFE are worth living.

Now we can ask: how much value must life itself have in order to satisfy the bad-life-worth-living intuition at the scale of human lives? Obviously, the answer will depend on how low one sets the average quality of BAD LIFE. But remember that it is stipulated that BAD LIFE is barely worth living by the lights of the positive theorist, meaning that its average quality is about as negative as it could be while still being outweighed by the value of life itself. If the positive theorist wishes to satisfy the bad-life-worth-living intuition at the scale of human lives, then the constant value of life itself must be large enough to outweigh the net badness accumulated over the course of a normal human life. To appreciate what that means, we need to think more concretely about the goods and bads within a normal human life.

The estimated global life expectancy for a person born in 2020 is 73 years.[14] Consider how many goods and bads are contained within 73 years of a normal human life. Take a moment to think about all of the goods and bads within your own life: from grade school all the way to the current point in your career, of all the experiences you have had with friends and family and teachers and students and bosses and colleagues and strangers, of your achievements and fortunes and defeats and hardships, and of little moments both meaningful and mundane. To appreciate how large the value of life must be if it is a constant (and if BAD LIFE is worth living), we need to shift from thinking about momentary experiences to the accumulation of goods and bads over the course of a whole lifetime. Even if the average quality of

---

[14] See United Nations, Dept. of Economic and Social Affairs, Population Division [2019].

BAD LIFE is only mildly negative, the value of life must be remarkably high in order to offset the accumulated badness from such a life over the course of 73 years. This means that if one thinks that the value of life itself is a constant, and if one wishes to satisfy the bad-life-worth-living intuition at the scale of human lives, then it will be hard to deny that life itself has a great amount of value.

At first, this appears to be a good result for the positive theorist. By taking the value of life to be a constant, the positive theorist can deny that ETERNITY is better than PARADISE while still accepting that BAD LIFE is worth living. The cost is that the positive theorist must qualify the bad-life-worth-living intuition by denying that it holds for lives of arbitrary length, but this also carries the benefit of justifying the claim that life itself has a great amount of value. In brief, it seems that we have found an escape from the Argument from Eternity. Unfortunately, the exit from ETERNITY leads to SWIFT HELL:

_____

SWIFT HELL: A horrible life that lasts for one minute.

_____

How horrible is SWIFT HELL? Let us stipulate that the value generated within SWIFT HELL is $-c + \epsilon$, where ϵ is some small number greater than zero. In other words, we define SWIFT HELL to be just about as awful as it can be while still having that badness outweighed by the goodness from life itself. Speaking somewhat metaphorically, we can think of SWIFT HELL as the result of taking all of the net badness from 73 years of BAD LIFE and condensing that into a single minute (and then tweaking it down just a bit). To put this into perspective, the average quality of SWIFT HELL is a little bit less than 38,368,800 times as negative as the average quality of BAD LIFE.[15] This suggests that even if the average quality of BAD LIFE is only mildly negative, SWIFT HELL must be unimaginably horrific.

_____

[15] Where does this number come from? SWIFT HELL lasts one minute, BAD LIFE lasts 73 years, and there are 525,600 minutes in a year. 525,600 × 73 = 38,368,800.

The Argument from Hell

P1: $c > 0$.

P2: The value generated by the bads within SWIFT HELL is $-c + \epsilon$ (where $\epsilon > 0$).

P3: The value generated by life itself for SWIFT HELL is c.

C: SWIFT HELL is a life worth living.

The reasoning is straightforward. By stipulation, the value generated within SWIFT HELL is $-c + \epsilon$, where ϵ is some small number above zero. By *Constant*, the value generated by life itself for SWIFT HELL is c. By the positive theory, c is greater than 0. Since any life with a value greater than zero is by definition a life worth living, it follows that SWIFT HELL is a life worth living. But by light of the reasoning from above, SWIFT HELL is an unimaginably horrific life. Therefore, the constant positive theory must accept that a very brief life full of nothing but horrific suffering is a life worth living. That is clearly the wrong result: SWIFT HELL is not worth living.

To resist the Argument from Hell, the positive theorist must reject one of its premises. P1 is the positive theory. Since P2 is a stipulation, it cannot be rejected. Therefore, the only option for the positive theorist is to reject P3, which is tantamount to rejecting *Constant*. The upshot is that both *Linear* and *Constant* lead to unacceptable results: the former leads to the Argument from Eternity while the latter leads to the Argument from Hell. This establishes the core of my argument against the positive theory. In the next section, I will continue developing the Argument for Zero by arguing that any principle besides *Linear* and *Constant* will still face some version of the objections I have advanced.

§ 3 | THE ARGUMENT AGAINST ASYMPTOTE

It is possible for a positive theorist to accept both that the value of life increases as a function of the length of life (as with *Linear*) and that the value of life has a maximal bound (as with *Constant*). The natural way of reconciling these claims is to hold that the value of life increases asymptotically as a function of the length of life, meaning that the marginal value from life itself grows smaller as the length of life grows larger:

Asymptote: The value of life increases asymptotically with the length of life.

As before, let c denote the maximal value of life itself. According to *Asymptote*, as a life grows arbitrarily long, the value generated by life itself approaches c. We also need to define a new constant, n, which specifies how quickly that maximal value is approached. More precisely, n marks how long a life must be in order to generate half of c:

c the maximal value of life itself
n the length of life that generates half of c

As an example, suppose we set $c = 100$ (meaning the maximal bound for the value of life itself is 100) and $n = 73$ (meaning that it takes 73 years for a life to generate half of the maximal value from life itself). Then the value generated by life itself converges to 100 as life gets arbitrarily long, with 73 years marking the length of time needed to generate 50 value from life itself.[16]

With these two constants, we can specify the equation for *Asymptote*: [17]

$$\lambda(L) = c \times \frac{\tau(L)}{\tau(L) + n}$$

The Asymptote Equation

This equation is slightly more complex, so let me simply point out a key observation. The higher the value of n, the more $\lambda(L)$ behaves like a linear function, while the lower the value of n, the more $\lambda(L)$ behaves like a constant function. In other words, as n approaches infinity, the asymptote equation approximates the linear equation, while as n approaches zero, the asymptote equation approximates the constant equation.

Asymptote enables the positive theorist to avoid both the Argument from Eternity (since there is a maximal bound for the value generated by life itself) and

---

[16] **Question to Reader:** Should I cut this paragraph with the example?

[17] There are other ways of constructing an asymptotic function, but this is the simplest form of an asymptotic function where positive inputs always yield positive outputs and that is monotonic (i.e., continuously increasing). Other asymptotic functions will still be susceptible to either the Argument from Eternity, the Argument from Hell, or the Argument against Asymptote developed in this section.

the Argument from Hell (since life itself would generate very little value after only one minute). Moreover, *Asymptote* satisfies both the bad-life-worth-living intuition and the intuition that more life means more value. These observations suggest that *Asymptote* provides an escape from the Argument for Zero. But unfortunately (for the positive theorist), the appearance of escape is illusory.

Here is the basic quandary. To avoid an analogue of the Argument from Eternity, one must ensure that *n* is not too high (so that the global value of ETERNITY in comparison to a normal human life is not too high). To avoid an analogue of the Argument from Hell, one must ensure that *n* is not too low (so that SWIFT HELL does not generate enough value from life itself for it to be worth living). But these constraints are in direct tension with each other. The question is whether there is some middle value of *n* that escapes unacceptable results in either direction. In what follows, I will argue that this quandary cannot be resolved by arguing that every value of *n* yields unacceptable results.

THE ARGUMENT AGAINST ASYMPTOTE

To develop the Argument against Asymptote, we need to define a new function, φ, which takes as input a life and outputs the value generated by the goods and bads within that life. Here again are all three of the functions this paper has defined:

$\tau(L)$ = length of L
$\lambda(L)$ = value of life itself for L
$\varphi(L)$ = value of goods and bads within L

We also need to define (or redefine) some lives. In the characterizations below, the first line provides an informal description of the life while the second line formally specifies the key properties of that life. As before, let ϵ be some very small number greater than zero:

——————————

ETERNITY: An empty life that lasts indefinitely.
              $\varphi(\text{ETERNITY}) = 0$ and $\lambda(\text{ETERNITY}) = c$

BAD LIFE: A bad life that lasts for 73 years.
              $\varphi(\text{BAD LIFE}) + \lambda(\text{BAD LIFE}) - \epsilon = 0$.

GOOD LIFE: A good life that lasts for 73 years.
              $\varphi(\text{GOOD LIFE}) = -\varphi(\text{BAD LIFE})$.

SHORT HELL: An awful life that lasts for 1 year.
              $\varphi(\text{SHORT HELL}) = -\lambda(\text{SHORT HELL}) + \epsilon$.

——————————

Let me briefly walk through these descriptions.

As before, ETERNITY is a life that contains no goods or bads but yields the maximal value from life itself. The formal gloss says that the value from within ETERNITY is zero while the value from life itself for ETERNITY is maximal.

As before, BAD LIFE is a life that is barely worth living by the lights of the positive theorist. The formal gloss says that the value from within BAD LIFE barely outweighs the value from life itself for BAD LIFE. In other words, the average quality of BAD life is stipulated to be almost as negative as it can be while still being outweighed by the value from life itself.

We can think of GOOD LIFE as a milder version of PARADISE. The formal gloss says that the goodness from within GOOD LIFE is exactly equal to the badness from within BAD LIFE. Note that while GOOD LIFE and BAD LIFE are symmetrical with respect to the goods and bads within them, they are asymmetrical with respect to global value. In particular, BAD LIFE is barely worth living according to the positive theorist (because the value generated by life itself barely outweighs the disvalue from within the life) while GOOD LIFE is very much worth living (because the value from life itself just adds further value on top of the value from within the life).

We can think of SHORT HELL as a stretched-out version of SWIFT HELL. The formal gloss says that the badness from within SHORT HELL is just barely outweighed

by the goodness from life itself for SHORT HELL. In other words, SHORT HELL is defined as being almost as awful as it could be while still being worth living.[18]

Now we are ready for the Argument against Asymptote. As a reminder, the asymptote equation involves a constant n, which specifies how long a life must last in order to generate half of the maximal value from life itself. In what follows, I will argue that no value of n avoids unacceptable results. My argument proceeds in three stages. First, I will argue that the greater the value of n (i.e., as n approaches infinity), the more unacceptable the results become with respect to ETERNITY. Second, I will argue that the lower the value of n (i.e., as n approaches zero), the more unacceptable the results become with respect to SHORT HELL. Third, I will argue that there is a middle value of n that yields unacceptable results in both directions. These arguments collectively comprise the Argument against Asymptote.

The Upper Limit

To begin, consider what happens as n approaches infinity. The larger the value of n, the smaller the ratio between (1) the value generated by a life of any particular finite length and (2) the maximal value of life itself. At the limit, when n = infinity, the maximal value from life itself will be infinitely times greater than the value from life itself for any finitely long life (making the ratio equal to zero). The bad-life-worth-living intuition says that a bad human life of 73 years is worth living. Because of this, we can think of the value of life itself at n = 73 as an anchor point. This means that the larger the value of n, the greater the value of ETERNITY:

$$\lim_{n \to \infty} \lambda(\text{ETERNITY}) = \infty$$

Value of ETERNITY as n approaches infinity

Here is the key point: as n grows larger, the asymptote equation begins to approximate the linear equation. The greater the value of n, the greater the value of ETERNITY. At the limit, when n is infinity, we run into the Argument from Eternity.

---

[18] It is possible to develop the Argument for Asymptote using PARADISE and SWIFT HELL (instead of GOOD LIFE and SHORT HELL). I choose to use GOOD LIFE and SHORT HELL because they make it somewhat easier to conceptualize the argument.

The Lower Limit

Next, consider what happens as n approaches zero. The smaller the value of n, the larger the ratio between (1) the value from life itself for a life of any particular finite length and (2) the maximal value of life itself. At the limit, when $n = 0$, the value generated by life itself for any finitely long life (even a very short one) will be equivalent to the maximal value from life itself (making the ratio equal to one). Now remember that $\varphi(\text{SHORT HELL}) = -\lambda(\text{SHORT HELL}) + \epsilon$. Therefore, the smaller the value of n, the worse SHORT HELL becomes:

$$\lim_{n \to 0} \varphi(\text{SHORT HELL}) = -c + \epsilon$$

Value within SHORT HELL as n approaches zero

Here is the key point: as n grows larger, the asymptote equation begins to approximate the constant equation. The smaller the value of n, the lower the quality of SHORT HELL. At the limit, when n is zero, we run into a version of the Argument from Hell.

A Middle Point

We have seen that *Asymptote* runs into problems if n is too high and if n is too low. The question now is whether there is some magic middle number that avoids unacceptable results in both directions. Suppose we set $n = 73$, as a somewhat arbitrary middle point between the limits of zero and infinity. If $n = 73$, then a life lasting 73 years would generate half of the maximal value that can be generated from life itself. My goal now is to show that if $n = 73$, then we get unacceptable results both with respect to ETERNITY and with respect to SHORT HELL. Given the conclusions of The Upper Limit and The Lower Limit, it will follow that no value of n avoids unacceptable results.

To develop this stage of the argument, it will be helpful to lay out the results for ETERNITY, BAD LIFE, GOOD LIFE, and SHORT HELL when we apply to those lives the functions for (1) length of life, (2) value from life itself (according to the asymptote equation with $n = 73$), and (3) value from within life. Here are those results:

Length of Life	Value from Life Itself	Value from within Life
$\tau(\text{ETERNITY}) = \infty$	$\lambda(\text{ETERNITY}) = c$	$\varphi(\text{ETERNITY}) = 0$
$\tau(\text{BAD LIFE}) = 73$	$\lambda(\text{BAD LIFE}) = \frac{1}{2}c$	$\varphi(\text{BAD LIFE}) = -\frac{1}{2}c + \epsilon$
$\tau(\text{GOOD LIFE}) = 73$	$\lambda(\text{GOOD LIFE}) = \frac{1}{2}c$	$\varphi(\text{GOOD LIFE}) = \frac{1}{2}c - \epsilon$
$\tau(\text{SHORT HELL}) = 1$	$\lambda(\text{SHORT HELL}) = \frac{1}{74}c$	$\varphi(\text{SHORT HELL}) = -\frac{1}{74}c + \epsilon$

These results, I will argue, are unacceptable.

Consider first ETERNITY. Recall that the global value of a life L is the sum of the value from within L and the value from L itself: formally, global value is $\lambda(\text{L}) + \varphi(\text{L})$. The results above show that the global value of ETERNITY is greater than the global value of GOOD LIFE. In particular, the global value of ETERNITY is c, whereas the global value of GOOD LIFE is $c - \epsilon$. This means if $n = 73$, then ETERNITY is better than GOOD LIFE. But that is implausible. If one had a choice about which life to live, it seems obvious that GOOD LIFE would be preferable to ETERNITY. Therefore, if $n = 73$, we have an unacceptable result concerning ETERNITY.

Consider next SHORT HELL. Recall that the average quality of a life L is the value within L divided by the length of L: formally, average quality is $\frac{\varphi(\text{L})}{\tau(\text{L})}$. The results above show that the average quality of SHORT HELL is nearly twice as negative as the average quality of BAD LIFE. In particular, the average quality of SHORT HELL is $-\frac{1}{74}c$ whereas the average quality of BAD LIFE is $-\frac{1}{146}c$. Yet the positive theorist must think that SHORT HELL is worth living, since we stipulated that the badness within SHORT HELL is outweighed by the value from life itself for SHORT HELL. This means that if $n = 73$, the positive theorist will be forced to say that a life much shorter and with a much worse average quality than BAD LIFE is still worth living. But that is implausible. How could a life that is both very bad and very short with no other redeeming qualities be worth living? Therefore, if $n = 73$, we have an unacceptable result concerning SHORT HELL.

There may be a temptation to argue that if the positive theorist thinks that BAD LIFE is still worth living, then they might likewise also think that a life with an average quality that is twice as bad as BAD LIFE is also still worth living. But remember that we stipulated that the average quality of BAD LIFE is as bad as it can possibly be while still being worth living by the lights of the positive theorist. In other words,

BAD LIFE was defined as a life that the positive theorist thinks is barely worth living. Yet whatever BAD LIFE looks like, SHORT HELL is approximately twice as bad with respect to average quality and much shorter with respect to length. This means that if the positive theorist accepts *Asymptote*, then for whatever life they initially think is barely worth living, there will be shorter lives with worse average qualities that are also worth living. I think nearly everyone will find this result unacceptable.

Therefore, $n = 73$ yields unacceptable results both with respect to ETERNITY and with respect to SHORT HELL.

The Argument Against Asymptote

The Argument against Asymptote has been complex, so let me provide a brief summary. Anyone who accepts *Asymptote* must specify a value for n, which marks how quickly the value of life itself approaches the maximal value of life itself. If $n = 73$, then we get the unacceptable results that (1) GOOD LIFE is worse than ETERNITY, and (2) SHORT HELL is worth living. If we raise the value of n, then the results become worse with respect to ETERNITY. If we lower the value of n, then the results become worse with respect to SHORT HELL. Therefore, no value of n avoids unacceptable results. So, we ought to reject *Asymptote*.

It is worth noting that my arguments do not require making any assumptions about what exactly BAD LIFE or GOOD LIFE or SHORT HELL look like. Nor do they require any specific view about how to quantify the value of any particular life. Instead, my arguments examine the structure of the most natural versions of the positive theory and play with that structure to expose implausible consequences of those theories. The source of the problem is the positive theory itself.

There are functions besides *Linear*, *Constant*, and *Asymptote* that behave in different ways, but I can think of no functions outside those classes that are promising candidates for developing a positive theory. In fact, there is reason to think that any function from the length of life to the value of life will be susceptible to some form of the Argument for Zero. For any such function, we can ask which length of life is *optimal*, meaning which length generates the maximal value from life itself. Then we can either (1) consider a life of optimal length whose average quality is negative but where the badness from within the life is outweighed by the value from life itself (as with SHORT HELL) or (2) compare an empty life of optimal length with a good life of a non-optimal length (as with ETERNITY and GOOD LIFE). I focused on

Linear, *Constant*, and *Asymptote* because those strike me as the best ways of developing the positive theory. But I suspect every positive theory that takes the value of life to be a function of the length of life will be vulnerable to the Argument for Zero.

Here is the upshot: *Linear* leads to the Argument from Eternity, *Constant* leads to the Argument from Hell, *Asymptote* leads to some version of either (or both) depending on the structure of the asymptotic function, and there seem to be no other promising principles for specifying how the value of life itself varies as a function of the length of life. This yields a more general version of the Argument for Zero that accommodates the results from the Argument against Asymptote:

The Argument for Zero

P1:   GOOD LIFE is better than ETERNITY.
P2:   SHORT HELL is not worth living.
P3:   If life is worth living for its own sake, then either P1 or P2 is false.
C: Life is not worth living for its own sake.

§ 4 | ARGUMENTS AGAINST ALTERNATIVES

In what follows, I will consider some other options for resisting the Argument for Zero. I will argue that none succeed.

Option 1: Rejecting Quantity

I have taken for granted the following principle:

Quantity: The value of life itself increases only as a function of the quantity of life.

I have also assumed that what it is to have a greater quantity of life is to have a greater length of life. Though there may be other ways of measuring quantity of life (such as in terms of number of experiences), it is straightforward to see how other quantification methods will still generate the Argument for Zero. But what if the positive theorist instead takes the value of life to increase as a function of some factor besides quantity of life? As a first pass, consider the following principle:

Diversity: The value of life increases as a function of its diversity of experiences.

According to *Diversity*, the value of life itself depends not on the quantity of life but instead on how rich and variegated the life is. ETERNITY contains only a single kind of experience, and SWIFT HELL contains nothing but horrible pain. These lives lack the texture and shape that characterize those of normal humans. By contrast, even though BAD LIFE has a negative average quality, it may still be rich and variegated enough to yield a large amount of value from life itself. However, consider the following life:

_____

DIVERSE HELL: A life filled with an extremely diverse set of experiences, all bad.

_____

Though DIVERSE HELL contains no good experiences, it still contains a rich variety of experiences: pain, nausea, hunger, thirst, fear, anger, anxiety, sadness, frustration, horror, disgust, itchiness, and so on. DIVERSE HELL is a life that involves suffering in as many ways as you can imagine and in many more ways you cannot imagine. If *Diversity* were true, then DIVERSE HELL may well be worth living, so long as the intensity of each experience is sufficiently mild. But DIVERSE HELL is not worth living, so we should reject *Diversity*.

Is there another way of rejecting *Quantity*? The form of argument developed above can be generalized. If the value of life itself is taken to be a function of some quantity, then we can always consider a life that scores high on that quantity yet is filled with bads. So long as that is possible, we will be able to generate analogues of the Argument for Zero. The only way to avoid this result would be to take the value of life to increase only as a function of the good (rather than the bad), as with the following principle:

Positivity: The value of life increases as a function of the number of goods within it.

Positivity predicts that both DIVERSE HELL and SHORT HELL are not worth living (since they contain no good experiences), that the value of ETERNITY is zero (since it contains no good experiences), and yet that BAD LIFE may nevertheless be worth

living (since it may still contain many good experiences). In fact, I think there is no simple counterexample to *Positivity*, for the principle is immune to the pattern of argument that all of the other principles have been vulnerable to. This is because *Positivity* is designed so that an increase in the bads within a life will always yield a net decrease in the global value of that life. But even if *Positivity* is immune to simple counterexamples, I think there is a deeper problem with the principle.

We began this paper with the following picture: life itself is good, meaning that any life acquires some goodness from life itself, regardless of the specific character of that life. That picture contrasts with the picture suggested by *Positivity*, where some lives generate no goodness at all, where the value of life itself is determined by (rather than independent of) the specific character of one's life, and where the amount of goodness from life itself dissociates from the quantity of life. That is a departure from the idea that life is valuable even when its goods are "too meager to outweigh" the bads because there is "additional positive weight" from life itself. In general, it is natural to think that if x is good, then more of x yields more goodness from x. By contrast, *Positivity* says that x is good, but that it is only more of y that yields more goodness from x. That is an odd axiological structure. It raises the question of why we should think that it is really life itself (rather than the having of more goods within life) that is good.

As an analogy, imagine a philosopher who claims that belief (rather than knowledge) is valuable while also claiming that the value of belief is defeated whenever a belief is false, unjustified, unsafe, or otherwise does not amount to knowledge. It would be natural to ask the following: if beliefs generate value when and only when they amount to knowledge, then why not think that it is simply knowledge that is valuable? Unless there is some independent and principled reason for thinking that it is belief itself that generates the value (and that the value of belief is defeated whenever it fails to amount to knowledge), we should instead favor the simpler hypothesis that it is knowledge that is valuable. Similarly, if one is sympathetic to the intuitions driving *Positivity*, then one should simply favor the hypothesis that diversity of good experiences is valuable.

This paper began with a metaphor, where we imagined a scale that weighs the good against the bad. The positive theory corresponded to a picture where the plate weighing the goods is heavier than the plate weighing the bads. But on the picture suggested by *Positivity*, the plates are perfectly balanced, at least before we

add any goods or bads to either side. On this picture, the plate weighing the goods itself becomes heavier only when goods are added to it. But if you favor that distribution of weight, why not simply think that the additional weight is carried by those goods themselves, rather than by the plate (but only when the goods are added)?

In light of this, I think that any theory that contends that the value of life itself is a function of some quantity besides the quantity of life carries a burden of explanation. The theory must explain why it is life itself that generates the value, rather than whatever quantity it is that serves at the scaling factor. Unless we have independent reason to believe that such a theory is true, we ought to think that the only candidates for positive theories are those that take the value of life itself to be a function of the quantity of life. If we grant that quantity of life is equivalent to length of life, then we arrive back at the Argument for Zero.

Option 2: Rejecting the Archimedean Property

I have taken for granted the following principle:

Archimedean Property: Let a be some value from within a life, and let b be the value from life itself. Then there is some constant c such that $a = c \times b$.

It is possible to reject the *Archimedean Property* by holding that some goods are superior to other goods, meaning that any amount of the former outweighs any amount of the latter.[19] More precisely, let us say that a good g_1 (or a bad b_1) is *superior* to a good g_2 just in case any amount of g_1 (or b_1) is better than (or outweighs) any amount of g_2. Now consider the following view:

Superiority: Every good or bad within life is superior to life itself.

---

[19] A number of philosophers have appealed to the idea that some goods are superior to other goods as a way of resisting the Repugnant Conclusion. See Arrhenius & Rabinowicz [2015], Parfit [2016], and Nebel [forthcoming] for some recent examples. Though I do not think superiority saves the positive theory, my arguments are compatible with superiority as a response to the Repugnant Conclusion.

This yields a new way for the positive theorist to resist the Argument for Zero: namely, by taking the goods and bads within life to be superior to the value of life itself. First, *Superiority* predicts that GOOD LIFE will be better than ETERNITY, since GOOD LIFE contains within it some goods while ETERNITY contains within it no goods. Second, *Superiority* predicts that SHORT HELL is not worth living, since the badness of the bads from within SHORT HELL is superior to the value from life itself.

The problem is that *Superiority* abandons the motivation that drives the positive theory in the first place: namely, the bad-life-worth-living intuition. If every good or bad within life is superior to life itself, then no version of BAD LIFE is worth living. In fact, as far as I can tell, there is only one respect in which the predictions of *Superiority* differ from the predictions of the neutral theory: namely, *Superiority* takes empty lives (i.e., lives that contain no goods nor bads) to be barely above the threshold of being worth living, while the neutral theory takes empty lives to be exactly on the threshold of being worth living. In any other case, the value of life itself may as well be zero. This is a far cry from the view we started with at the beginning of this paper.

Could the positive theorist adopt a weaker version of *Superiority*? A first approach is to hold that for any good (or bad) within life, *some* (rather than *any*) amount of that good (or bad) is better than (or outweighs) any amount of value from life itself. A second approach is to hold that some goods and bads are not superior to life itself. Either of these weakenings would enable the positive theorist to retain the bad-life-worth-living intuition. But the problem is that these views no longer render the positive theory immune to the Arguments for Zero. If some amounts of some goods are not superior to life itself, then we can consider a version of GOOD LIFE whose quantity of goods does not exceed that threshold in order to generate the result that ETERNITY is better than GOOD LIFE. If some amounts of some bads are not superior to life itself, then we can consider a version of SHORT HELL whose quantity of bads does not exceed that threshold in order to generate the result that SHORT HELL is worth living. As soon as the positive theorist weakens *Superiority* to save the bad-life-worth-living intuition, the Argument for Zero slips through the cracks.

Option 3: Rejecting Globalism

I have assumed the following:

Globalism: Life A is better than life B just in case the global value of A is greater than the global value of B.

Though Globalism is an extremely popular principle, perhaps the positive theorist could reject it. A first approach is to appeal to average values rather than global values. Recall that the average value of a life is the global value of that life divided by the length of that life. Let α be a function from lives to average values. Then we have the following view:

Average Value: A is better than B just in case $\alpha(A) > \alpha(B)$.

Average Value is a principle that determines what makes one life better than another, rather than a principle concerning the value of life itself. But the positive theorist can simply adopt any of the prior principles relating the value of life itself to the length of life. Suppose, for example, that the positive theorist adopts *Linear*. Then we can reassess our comparisons of the lives we have defined. The average value of ETERNITY is negligible, the average value of GOOD LIFE is positive, and the average value of BAD LIFE is barely positive. This means that *Average Value* predicts that GOOD LIFE is better than ETERNITY and that BAD LIFE is worth living. There may be a temptation to think that *Average Value* runs into a problem with SHORT HELL. But given the way that SHORT HELL was defined, and supposing we favor *Linear*, it will turn out that the average quality of SHORT HELL is no worse than the average quality of BAD LIFE. Therefore, *Average Value + Linear* provides a way out of the dilemma generated by the Argument for Zero.

The problem is that *Average Value* transforms the dilemma into a trilemma. Consider the following life:

_____

SWIFT HEAVEN: A blissful life that lasts for one minute.

_____

Let us stipulate that the average value of SWIFT HEAVEN is slightly higher than the average value of PARADISE. Then *Average Value* predicts that SWIFT HEAVEN

is better than PARADISE (which has a slightly worse average value but which lasts for 73 years). But PARADISE is obviously better. Therefore, we ought to reject *Average Value*.[20]

A second approach is to hold that the longer a life lasts, the less weight that is assigned to any individual point within that life. A sophisticated version of this is developed by…

Are there other ways of rejecting *Globalism*? I do not know of any other approaches that are promising. In fact, I think it is no coincidence that *Globalism* (and its interpersonal analogue) is often taken for granted in contemporary population ethics. Despite the fact that *Globalism* generates challenges (such as the Repugnant Conclusion), its appeal is considerable and the alternatives are questionable.

Option 4: Rejecting the bad-life-worth-living intuition

Could the positive theorist simply forfeit the bad-life-worth-living intuition?

Any positive theory must accept that some version of BAD LIFE is worth living. Even if we allow the value of life itself to be arbitrarily small, we could consider versions of BAD LIFE whose average qualities are arbitrarily close to zero and whose lengths are arbitrarily close to zero. But this means that the positive theorist is still left with the same results as before: any version of *Linear* will still be faced with the result that some life whose average quality is as good as GOOD LIFE is worse than ETERNITY, any version of *Constant* will be faced with the result that some very short lives involving nothing but pain will still be worth living, and any version of *Asymptote* will be faced with both of the aforementioned results.

The positive theorist might point out that the results are less counterintuitive when we take the value of life itself to be sufficiently small. For example, if the value of life itself is extremely small, then ETERNITY would be better than only a very brief version of GOOD LIFE and only mild versions of SHORT HELL would be worth living. However, softening the blow of the counterintuitive results carries the cost of

---

[20] The challenges for average welfare principles are familiar from prior discussions in population ethics, such as Parfit [1984, p. 420] and Huemer [2008]. For a recent defense of average utilitarianism, see Pressman [2015]. Notably, Pressman stresses that average utilitarianism (about populations of lives) is compatible with globalism (about individual lives). To my knowledge, nobody has explicitly argued for *Average Value* about individual lives.

sapping the interest from the positive theory itself. We began this paper with a philosophically stimulating picture, according to which even a life where the bads are plentiful and the goods are scarce would be worth living because of the extra goodness from life itself. That picture gradually dissolves as the value of life itself gradually shrinks.

Out of Options?

In §2, I presented the Argument for Zero. In §3, I argued that every version of the positive theory that takes the value of life to be a function of the length of life faces some version of the Argument for Zero. In §4, I argued against approaches that involve rejecting *Quantity*, the *Archimedean Property*, *Globalism*, or the bad-life-worth-living intuition. I cannot think of any other defensible ways of resisting the Argument for Zero. Therefore, I conclude that the positive theory is false. Supposing that the negative theory is a non-starter, the only option is the neutral theory. This concludes my argument that life is neutral.

§ 5 | PHILOSOPHICAL CONSEQUENCES

This final section briefly discusses the question of whether life itself is good relates to other philosophical questions about the threshold for a life worth living, the basic welfare goods, morality and rights, the value of consciousness, the non-identity problem, a variety of issues in applied ethics concerning life and death, and whether death is bad.

The first and most obvious connection concerns the notion of a life worth living. If the positive theory were correct, then some lives with negative average qualities would lie above the threshold for a life worth living. By contrast, on the neutral theory, whether or not a life is worth living depends only on the goods and bads within that life. A noteworthy observation is that the positive theory would render the Repugnant Conclusion especially repugnant, for it would entail that for any finite population A containing only lives with excellent average qualities, there is a larger population Z containing only lives with negative average qualities such that Z is better than A. By contrast, the neutral theory requires that a life barely worth living still have a positive average quality, tempering the repugnance of the Repugnant Conclusion.

A second connection concerns questions about the basic welfare goods. Philosophical discussions of welfare goods are mostly centered on debates between experientialists, desire-satisfactionists, and objective-list theorists. But on occasion, life itself shows up on lists of the basic welfare goods, and philosophers sometimes simply claim that life itself is good.[21] In fact, life itself is especially interesting as a candidate for a welfare good, since we are asking whether that which contains welfare goods is itself a good. By arguing for the neutral theory, I have argued that the only viable candidates for welfare goods are those contained within lives.

This result about welfare goods bears on the common idea that anything that has a life (in the sense of being a welfare subject) thereby has moral status.[22] If the positive theory were correct, then one way to justify that claim would be to appeal to (1) the fact that anything that has a life benefits from some welfare good (namely, life itself), and (2) the idea that anything that benefits from some welfare good has moral status. However, this explanation is unavailable to the neutral theorist. This means that a neutral theorist who accepts that anything that has a life thereby has moral status must find a different justification for that claim.

It may be tempting to think that the neutral theory also has bearing on the idea that there is a right to life (meaning that individuals have a right to not be killed by other agents). If life itself is not good, then why think that there is a right to life? However, I think the neutral theory is compatible with thinking that there exists a right to life. It may be that while life is not good, one still has a right to life because life is what enables one to acquire goods. Compare this to a view that says that freedom itself is not good, but that one still has a right to freedom because freedom is what enables one to acquire goods.

A third connection concerns a parallel debate about the value of consciousness.[23] The question of whether life itself is good is structurally analogous to the question of whether consciousness is intrinsically valuable. In fact, a common idea is that what it is to have a life is to be conscious.[24] The passage from Nagel quoted

---

[21] See, for example, Frankena [1973]'s list of candidates for welfare goods. See Campbell [2016] and Crisp [2017] for overviews on well-being.

[22] See Warren [1997] and Jaworska & Tannenbaum [2018] for overviews of moral status.

[23] See Lee [2018] and Kriegel [2019] for recent discussion.

[24] See, for example, Sumner [1996, p. 43], Kahane & Savulescu [2009], and Rosati [2009].

at the beginning of this paper even ascribed the goodness to "experience itself." Therefore, unless there is some reason to think that life and consciousness are disanalogous in some relevant way, the Argument for Zero indicates that consciousness is neutral.

By contrast, the question of whether life itself is good ought to be distinguished from debates about the value of existence. This is because debates about the value of existence are focused on whether some outcome can be better or worse for one than not existing at all. These debates largely stem from ethical puzzles such as non-identity problem, or the question of whether some acts that bring individuals into existence are wrong even though they are not bad for anyone. Since these kinds of puzzles about existence arise whether we accept the positive theory or the neutral theory, debates about the value of existence are largely independent from our question about the value of life.[25]

A fourth connection concerns a cluster of issues in the applied ethics of life and death, including issues about euthanasia, abortion, suicide, procreation, vegetarianism, and murder. In brief, the neutral theory suggests that we ought not preserve a person's life simply for the sake of enabling them to live longer, that depriving a fetus of life itself is not a reason against abortion, that the loss of life itself is not a reason against suicide, that the giving of life itself is not a reason for procreation, that the justification for vegetarianism ought to focus on the harms conferred to animals rather than merely loss of life, and that the wrongness of murder is not grounded in depriving the victim of life itself. There is plenty more to say about each of these issues, but I hope these brief observations illustrate why the question of whether life is worth living for its own sake has a range of practical implications.

Finally, the neutral theory has implications for how we think about the harm of death. The standard theory of the harm of death says that death is harmful in virtue of depriving an individual of goods that they would have had if they had not died. If the positive theory were correct, then every death would be at least somewhat harmful, since death necessarily deprives one of the good of life itself. But on

---

[25] There is, of course, a distinct question of whether existence itself is good. That parallels the question of whether life itself is good (and whether consciousness is intrinsically valuable). My point is simply that fall under the label 'value of existence' within the contemporary philosophical literature are independent of the issues discussed in this paper.

the neutral theory, it may be possible for there to be deaths that are not at all harmful, since what death deprives one of depends on contingent facts about which goods and bads one's life would have contained had one continued to live. In this respect, just as life is neutral, so too is death.[26]

CONCLUSION

This paper began with a picture where life is worth living for its own sake: the goodness of a life is determined not only by the goods within the life but also by the fact that life itself is good. Over several stages, I presented the Argument for Zero. The core argument was that taking life itself to be good leads either to the Argument from Eternity, or the Argument from Hell, or some more moderate (though still unacceptable) version of either. Therefore, every version of the view that life is worth living for its own sake generates unacceptable results. This leads us to the neutral theory, according to which life itself is neither good nor bad. According to the neutral theory, to determine how good a life is, or whether it is worth living, or whether it is better than another life, we need to look at the character of that life (rather than at the quantity of life). The goodness of a life is determined only by the goods within a life, for life itself is neutral.

---

[26] See Nagel [1970] for a classic article developing this view, and Luper [2019] for a recent overview of the philosophy of death.

REFERENCES

Agar, N., 2001. Life's Intrinsic Value, New York: Columbia University Press.

Arrhenius, Gustaf & Rabinowicz, Wlodek (2015). Value Superiority. In Iwao Hirose & Jonas Olson (eds.), *The Oxford Handbook of Value Theory*. New York, USA: Oxford University Press USA. pp. 225-248.

Arrhenius, Gustaf, Ryberg, Jesper and Tännsjö, Torbjörn, "The Repugnant Conclusion", *The Stanford Encyclopedia of Philosophy* (Spring 2017 Edition), Edward N. Zalta (ed.), URL = <https://plato.stanford.edu/archives/spr2017/entries/repugnant-conclusion/>.

Campbell, Stephen M. (2016). The Concept of Well-Being. In Guy Fletcher (ed.), Routledge Handbook of Philosophy of Well-Being. Routledge.

Crisp, Roger, "Well-Being", *The Stanford Encyclopedia of Philosophy* (Fall 2017 Edition), Edward N. Zalta (ed.), URL = <https://plato.stanford.edu/archives/fall2017/entries/well-being/>.

Frankena, William K., 1973, *Ethics*, second edition, Englewood Cliffs: Prentice Hall. [1, 6]

Glannon, Walter (2016). The Value and Disvalue of Consciousness. Cambridge Quarterly of Healthcare Ethics 25 (4):600-612.

Griffin, James (1986). Well-Being: Its Meaning, Measurement and Moral Importance. Clarendon Press.

Huemer, Michael (2008). In defence of repugnance. *Mind* 117 (468):899-933.

Jaworska, Agnieszka and Tannenbaum, Julie, "The Grounds of Moral Status", The Stanford Encyclopedia of Philosophy (Spring 2018 Edition), Edward N. Zalta (ed.), URL = <https://plato.stanford.edu/archives/spr2018/entries/grounds-moral-status/>.

Kriegel, Uriah (2019). The Value of Consciousness. Analysis 79 (3):503-520.

Lamb, David (1988). *Down the Slippery Slope: Arguing in Applied Ethics*. Routledge.

Lee, Andrew Y. (2018). Is consciousness intrinsically valuable? *Philosophical Studies* 176 (3):655-671.

Link, Hans-Jürgen (2013). Playing God and the Intrinsic Value of Life: Moral Problems for Synthetic Biology? *Science and Engineering Ethics* 19 (2):435-448.

Luper, Steven, "Death", *The Stanford Encyclopedia of Philosophy* (Winter 2019 Edition), Edward N. Zalta (ed.), URL = <https://plato.stanford.edu/archives/win2019/entries/death/>.

Nagel, Thomas (1970). Death. *Noûs* 4 (1):73-80.

Nagel, Thomas (2012). *Mortal Questions*. Cambridge University Press.

Nebel, Jake (forthcoming). Totalism Without Repugnance. In *Festschrift for Derek Parfit*.

Parfit, Derek (1984). *Reasons and Persons*. Oxford University Press.

Parfit, Derek (2016). Can We Avoid the Repugnant Conclusion? *Theoria* 82 (2):110-127.

Pressman, Michael (2015). A Defence of Average Utilitarianism. *Utilitas* 27 (4):389-424.

Schumacher, Bernard N. (2010). *Death and Mortality in Contemporary Philosophy*. Cambridge University Press.

United Nations, Department of Economic and Social Affairs, Population Division (2019). World Population Prospects 2019: Highlights. ST/ESA/SER.A/423.

Warren, Mary Anne (1997). Moral Status: Obligations to Persons and Other Living Things. Clarendon Press.

Chang, Ruth (2004). Putting together morality and well-being. In Peter Baumann & Monika Betzler (eds.), Practical Conflicts: New Philosophical Essays. Cambridge: Cambridge University Press. pp. 118--158.

Kagan, Shelly (2012). Death. Yale University Press.

Kagan, Shelly (2019). How to Count Animals, More or Less. Oxford University Press.