University of Kentucky

From the SelectedWorks of James M. Donovan

August, 2000

Reconciling Arguments Regarding a Self-Preserving Soul with Evolutionary Theory

James M Donovan

Available at: https://works.bepress.com/james_donovan/45/

RECONCILING ARGUMENTS ABOUT THE SELF-PRESERVING SOUL WITH EVOLUTIONARY THEORY

A Thesis

Submitted to the Graduate Faculty of the
Louisiana State University and
Agricultural and Mechanical College
in partial fulfillment of the
requirements for the degree of
Master of Arts

in

The Department of Philosophy

by

James M. Donovan
B.A., University of Tennessee at Chattanooga, 1981
M.L.I.S., Louisiana State University, 1989
Ph.D., Tulane University, 1994
August 2000

ACKNOWLEDGMENTS

The present work is a natural outcome of my long ago discussions with Andy Walker and John Miller. I continue to be indebted for their stimulating challenge. The death of my partner Jorge Vasconez raised the intellectual questions of this thesis to an even greater urgency. Religious people tried to console me by assuring me that "we would meet again." I did not expect to determine whether they were right, but I did need to know if the belief was reasonable. I wish I could have concluded affirmatively; I will miss him. My present life has been made full again by Gary Simmons, an unlooked for and unmerited blessing, the irony of which does not escape me.

TABLE OF CONTENTS

ACKNOWLEDGMENTS ii

ABSTRACT. v

PART I: THE PROBLEM BACKGROUND 1

CHAPTER I: RELATING SOULS TO BODIES 2
- A. INTRODUCTION. . . . . .
  - 1. The problem generally stated. . . . . .
  - 2. Philosophical assumptions . . . . . .
- B. A METHOD TO COMPARE SOUL-THEORIES WITH EVOLUTIONARY THEORIES. . . . . .
  - 1. Popperian Verisimilitude . . . . . .
  - 2. Potential weaknesses of this method . . . . . .
- C. POSSIBLE ARGUMENTS TO AVOID THE CONFLICT OF THEORIES . . . . . .
  - 1. Religious stances are exempt from empirical Criticism. . . . . .
  - 2. "Humans" versus "*Homo sapiens*" . . . . . .
  - 3. Argument *ex consensu gentium* . . . . . .
- D. FUNDAMENTALS OF EVOLUTIONARY THEORY . . . . . .
  - 1. Evolution, natural selection, and Darwinism. . . . . .
  - 2. Gradualism versus punctuated equilibria . . . . . .
- E. THE SELF PRESERVED BY THE SOUL. . . . . .
- F. THE CRITERION FOR COMPATIBILITY WITH EVOLUTION . . . .
- G. PROJECT OUTLINE. . . . . .

PART II: PHILOSOPHICAL TEXTS. . . . . .

CHAPTER II: PLATO'S *PHAEDO* (c. 387-380 B.C.E.). . . . . .
- A. CRITERION ONE: THEORY CONSISTENCY. . . . . .
  - 1. The opposition argument (69e-72d). . . . . .
  - 2. The recollection argument (72e-77a) . . . . . .
  - 3. The formal argument (78b-80c) . . . . . .
  - 4. The entailment argument (102a-107b). . . . . .
- B. CRITERION TWO: KIND OF SOUL. . . . . .
- C. CRITERION THREE: PSYCHOLOGICAL CONTINUITY . . . . . .
- D. CONCLUSIONS. . . . . .

CHAPTER III: ARISTOTLE'S *DE ANIMA* (C. 325 B.C.E.). . . . . .
A. CRITERION ONE: THEORY CONSISTENCY . . . . . .
B. CRITERION TWO: KIND OF SOUL . . . . . .
C. CRITERION THREE: PSYCHOLOGICAL CONTINUITY. . . . . .
D. CONCLUSIONS . . . . . .

CHAPTER IV: WILLIAM OF AUVERGNE'S *THE IMMORTALITY OF THE SOUL* (C. 1128-35) . . . . . .
A. CRITERION ONE: THEORY CONSISTENCY. . . . . .
1. The argument from dependence . . . . . .
2. The argument from incorruptibility. . . . . .
3. The argument from natural desires. . . . . .
4. The argument from destructibility. . . . . .
5. The argument from contrast with the senses. . . . . .
B. CRITERION TWO: KIND OF SOUL. . . . . .
C. CRITERION THREE: PSYCHOLOGICAL CONTINUITY . . . . . .
D. CONCLUSIONS. . . . . .

CHAPTER V: DAVID HUME'S *OF THE IMMORTALITY OF THE SOUL* (1777). . . . . .
A. CRITERION ONE: THEORY CONSISTENCY. . . . . .
1. Metaphysical arguments. . . . . .
2. Physical arguments. . . . . .
B. CRITERION TWO: KIND OF SOUL. . . . . .
C. CRITERION THREE: PSYCHOLOGICAL CONTINUITY . . . . . .
D. CONCLUSIONS. . . . . .

CHAPTER VI: MARTIN HEIDEGGER'S *BEING AND TIME* (1927) . . . . . .
A. CRITERION ONE: THEORY CONSISTENCY . . . . . .
B. CRITERION TWO: KIND OF SOUL . . . . . .
C. CRITERION THREE: PSYCHOLOGICAL CONTINUITY. . . . . .
D. CONCLUSIONS . . . . . .

PART III: CONCLUSIONS. . . . . .

CHAPTER VII: SHOULD WE BELIEVE OR NOT?. . . . . .

BIBLIOGRAPHY. . . . . .

VITA. . . . . .

ABSTRACT

Advocates of a self-preserving soul must somehow integrate that soul with an organic body which is the product of an evolutionary history. Efforts to resolve this version of the sorites paradox (at what point in the gradual evolution of *Homo sapiens* was the step taken from the soul-less to the ensouled?) can result in conflict between the two theories. The position is taken here that, given the standard of Popperian verisimilitude, any such conflict signals a weakness in the soul-theory, and not in evolutionism.

Five soul-theories (those by Plato, Aristotle, William of Auvergne, Hume, and Heidegger) are evaluated on three criteria. First, is the soul-theory internally consistent, but within itself and relative to the larger corpus of that author? Second, is the kind of soul depicted by that theory one which the ordinary person would recognize as his or her "self"? And third, is the soul-theory compatible with evolutionism? The standard for evolutionary compatibility is whether the theory allows for psychological continuity across species.

No theory successfully addresses all three criteria. Of the five, Hume's theory seems most promising. Had the evolutionary standard not been included, the better choice would have been Heidegger's. This fact demonstrates that evolutionary considerations do have philosophical consequences.

PART I

THE PROBLEM BACKGROUND

CHAPTER I

RELATING SOULS TO BODIES

A. Introduction

1. The problem generally stated

This thesis examines the relationship between implications of evolutionary biology and a concept of soul which preserves the self beyond the event of physical death. That a systematic problem exists is signaled by the inherently conflicting assumptions about development contained in each perspective. Simply phrased, bodies develop, souls don't, and somehow the one must be grafted onto the other.

On the one hand, evolutionary biology claims that the sort of creature we are is the end result of a long, piecemeal process of historical development. Debates still continue about exact sequences or particular mechanisms and rates for specific developments. But all in all, a general bottom-line summary of the theory of evolution (to be qualified below) would be: genealogical continuity between species reflecting gradual changes over time in natural reproductive populations. Evolutionary entities develop, meaning that they emerge over time.

On the other hand, nearly all conceptions of the soul make it out to be a thing *in toto*, which either is or is not, admitting of no gradations and which is unique to human beings. One may perhaps speak of a soul as being less or more advanced, but not as being less or more of a soul.

Passages which might seem to be speaking in these terms[1] are better read as describing the condition of the soul, or the differential development of powers within the soul, not about degrees of soul *per se*. There is no history of coming-into-being for the soul, either in the species or in the person. Whether or not it is immortal, it is generally presumed to be eternal, complete at every instant that it exists.

The intellectual problem is therefore this: biological beings are argued to possess immaterial souls. How are ahistorical souls to be overlain upon historical organisms, given that the developmental gradualism inherent in evolutionary biology conflicts with the categorical quality of the soul?[2] Yet the temporal and the eternal must be reconciled by any who would meaningfully assert that humans have souls.

Since we know that we became humans gradually, either of two claims must be made about the acquisition of the soul. First, perhaps we became possessed of the soul all at once. This seems to be the point

---

[1]Suggested passages of this kind have included those from Plato's *Phaedo* describing souls becoming "patched with corporeality," and the picture of a modular soul in Aristotle's *Nichomachean Ethics*.

[2]This thesis takes soul-theorists at their word that the soul exists, and does not seek to examine this question apart from critiquing their own attempts to rationally defend the assertion. A failed defense is therefore construed as a failure to justify the belief, and not as evidence that the belief is false. This latter conclusion is a different kind of argument than the one being made here, and to include it would force the thesis into many digressions. This position is not unprecedented: My love for a specific man is similarly both real and unjustifiable; it merely is, as a brute fact.

of the Roman Catholic position that "God created man…and to most interpreters this has implied the injection of a soul into a seminal population of our species."[3] This would remove the point of immediate conflict with evolutionary thinking, but only by increasing the demands upon the soul theory itself. This solution forces the spectre of the first (completely) ensouled child being born of (utterly) soul-less parents. If this sudden spiritual advance was arbitrary, it could just as easily have been accomplished by some other creature. There would thus still remain for the soul-theorist the question why *we* have souls and rocks do not, with the limitation that the answer cannot be phrased in terms of the physical differences (which it is the point of this strategy to avoid considering) *or* of the metaphysical differences, since it is the goal of this explanation to explain these, and hence they cannot be used as part of that explanation. The outcome, then, is that if the soul-theorist wishes to assert that the soul made an instantaneous, arbitrary appearance in our species, he cannot explain that fact either physically or metaphysically, which is to say he cannot explain it at all. Alternatively, the sudden appearance of the soul could be claimed to be linked to some identifiable features, anatomical or otherwise, which made this first ensouled child a suitable vehicle for this new construct.

---

[3]Steven M. Stanley, *The New Evolutionary Timetable* (New York: Basic Books, 1981), 167.

The question here becomes, What property does the soul supervene upon? What marks the difference between the soul-less parent and the ensouled child?

This question can be phrased as a form of the sorites paradox. The term refers to the Greek word for "heap," and asks the question, how many grains of wheat make up a heap? Surely not one, or even two, but almost certainly one thousand. At what point does the change come? In principle

> no one grain of wheat can be identified as making the difference between being a heap and not being a heap. Given then that one grain of wheat does not make a heap, it would seem to follow that two do not, thus three do not, and so on. In the end it would appear that no amount of wheat can make a heap," which is obviously false.[4]

---

[4]Dominic Hyde, "Sorites Paradox," *The Stanford Encyclopedia of Philosophy* (Spring 2000 Edition), Edward N. Zalta (ed.), URL = http://plato.stanford.edu/ archives/spring2000/entries/sorities-paradox.

My own impression is that in this form the paradox is more apparent than real. A heap is more reliably identified by the shape or configuration of its constituents, than by the number of those constituents. A thousand grains in a line is not a heap, while the same thousand in a pile is. So also I think very few grains can be a heap if they are properly arranged. There are, of course, other forms of the paradox which are not amenable to this resolution. Thomson, for instance, raises the question about where the line is to be drawn between the right-less fertilized egg and the right-holding baby. Judith Jarvis Thomson, *The Realm of Rights* (Cambridge, MA: Harvard University Press, 1990), 292. Derek Parfit uses a form of the sorites paradox to devastating effect in his argument that there is no "self." Derek Parfit *Reasons and Persons* (Oxford: Clarendon Press, 1984), 231-243.

In the present context the paradox becomes: How much of X renders an organism a likely possessor of a soul, X being variously defined? For any X (e.g., rationality), there is some little amount which most would agree is insufficient to support the soul (the amount found in animals, for example), and some greater amount which justifies that conclusion. But at what point is the line crossed?

This is the same question which must be addressed by the rare soul-theorist who believes that the soul did not appear instantaneously, but rather gradually like the body that possess it. Here, too, we must ask what are the evolutionary precursors of the soul, and what changes resulted in its mature appearance?

Either approach to the soul — that it appeared suddenly or gradually — requires some kind of confrontation with the fact of biological evolution. The only way to escape this confrontation is to eschew the project of rationally explaining the soul at all, as one does when one claims that the appearance of the soul was instantaneous and arbitrary.

2. Philosophical assumptions

The possibility of examining this topic depends upon the acceptance of two philosophical assumptions. Beyond the few lines given below these initial premises are not here defended. They are instead

taken as part of the problem background generating the present work which should be brought to the attention of the reader.

First, an external reality exists, independent of the person; second, the human being that we are is capable of gaining nontrivial knowledge about that external reality.[5]

Both of these tenets are denied by the postmodern philosophical movement.[6] Postmodernists would variously argue (a) that to the extent that "truth" is a correspondence between our linguistic statements about reality and the actual state of that reality,[7] there is no "truth" because

---

[5]Unknown to me at the time, George Gaylord Simpson identified this same "basic philosophical position" when prefacing his discussion on the implications of evolution.

> It is assumed that a material universe exists and that it corresponds with our perceptions of it. The existence of absolute, objective truth is taken for granted as well as the approximation to this truth of the results of repeated observations and experiments.

George Gaylord Simpson, *The Meaning of Evolution: A Study of the History of Life and of its Significance for Man* (New Haven, CT: Yale University Press, 1950), 7.

[6]For a critical appraisal of postmodernism as an intellectual school, see P.R. Gross and N. Levitt, *Higher Superstition* (Baltimore, MD: Johns Hopkins University Press, 1994).

[7]Not all philosophers, of course, interpret truth as a correspondence relationship. Martin Heidegger, for example, considers this interpretation to be derivative of a more fundamental understanding of truth as unconcealment. Karl Popper's position, however, is the one largely guiding this thesis, and he considers Tarski's "rehabilitation of the correspondence theory of truth" as one of the most significant

there is no external reality to which our statements can correspond.[8] Pulsars are for postmodernists socially created by astronomers, not discovered by them.[9] No statement about reality is therefore privileged relative to any other. From this view, "science" is different from, but no better than, "mythology," and is often held to be worse because it is the product of an oppressive white male worldview bent on domination and control of both nature and other people; or (b) that even were there an external reality which exists independently of us, we have no reliable access to it.

---

accomplishments of his time. Karl Popper, *Objective Knowledge* (London: Oxford at the Clarendon Press, 1972), 328.

> There is no doubt that correspondence to the facts is what we usually call 'truth'; that in ordinary language it is correspondence that we call 'truth', rather than coherence or pragmatic usefulness. A judge who admonishes a witness to speak the truth and nothing but the truth does not admonish a witness to speak what he thinks is useful either for himself or for anybody else. The judge admonishes a witness to speak the truth and nothing but the truth, but he does not say, 'All we require of you is that you do not get involved in contradictions', which he would say were he a believer in the coherence theory. (pp. 316-317)

[8]Popper cogently observes that falsifications of hypotheses largely disprove this belief. "They teach us the unexpected; and they reassure us that, although our theories are made by ourselves, although they are our own inventions, they are none the less genuine assertions about the world; for they can *clash* with something we never made." *Objective Knowledge*, 196-198.

[9]P.B. Roscoe, "The Perils of Positivism in Cultural Anthropology," *American Anthropologist*, 97(1997), 492-504.

Here the problem is that our perceptions are guided (or tainted) by our presuppositions. This influence need not be characterized as malignant. Karl Popper (more from whom below) asserted that "the growth of all knowledge consists in the modification of previous knowledge,"[10] and therefore that knowledge is "dispositional and expectational."[11] The postmodern twist is to regard these dispositions and expectations not as way of *relating* to the external world, but as wholly *constituting* that picture we label the external world. Whereas Popper might agree that we usually see what we expect to see (and hear, taste, touch, and smell; the senses are, in his terms, "theory-impregnated"), these are default inclinations, and not determinative. On the other hand, the postmodernist would argue that this is all we *can* see, hear, etc. We can see only what we expect to see because that is all that our minds are prepared to recognize.

If that were true, then we have no access to any external reality that we do not already expect to encounter. Truth as correspondence fails because our statements (which carry the property of being true or

---

[10]Popper, *Objective Knowledge*, 71.

[11]Ibid. Stephen Jay Gould echoes this position: "Information always reaches us through the strong filters of culture, hope, and expectation." Stephen Jay Gould, *The Panda's Thumb* (New York: W.W. Norton, 1980), 118. For this reason pure inductionism, the atheoretical, neutral accumulation of facts until a pattern emerges, does not work since what is recognized as a "fact" in the first place is established *a priori* by theory.

false) refer not to the external world but to other of our own mental states. Because no knowledge is true, it must be evaluated by some other standard (utility, perhaps, or whether it promotes self-esteem[12]) which will vary according to the relative needs of the knower. Again the result is extreme relativism, which is incompatible with the present exercise.

Objections not withstanding, these two initial tenets combine to form the following standard: *There being only one external reality, any two true statements about the same thing must not be in irresolvable conflict.* Incompatibilities due to our relative lack of knowledge are tolerable, at least for a short time, and these are scientific problems. The philosophical problem arises when statements are incompatible in principle, such that no new knowledge could be expected to ameliorate the conflict.

When this happens, the proper response is not, as might a postmodern social scientist, to opine that the two statements are merely different "perspectives" on the same thing, and thus both deserving of respect. Rather, at least one statement must not be true. It is possible

---

[12]One university graduate student, for example, on the occasion of a series of campus lectures on the Darwinian legacy, admitted that "I don't agree with Darwin, because [his theory] undercuts the inherent dignity of human beings." S. Majoria, "Speaker Lectures about Darwinian Legacy to the Scientific Community," *The Reveille [Louisiana State University]*, October 6, 1998, 3+.

that both statements are false (if they are contraries and not contradictories), but it cannot be the case that two conflicting statements are both true if they are genuinely "about the same thing."

Given that standard, we may assert the following. If evolutionary and soul theories are in irresolvable conflict, then one at least must be false.

B. A Method to Compare Soul-Theories with Evolutionary Theory

1. Popperian verisimilitude

The soul-theorist could well accept what has been thus far argued: that the rational defense of a self-preserving soul requires consideration of the implications of organic evolution, and that any irresolvable conflict emerging from this comparison necessarily means that one of the two theories is false. He need not likewise accept the implication that the theory which is more probably false is his own. Conceivably he might assert the incompatability as a refutation of evolution.

Indeed, many soul-advocates have denied the validity of evolutionary theory altogether. Traditionally, our mental and moral faculties were seen as the product of a distinct spiritual agency, the "soul," added to the physical body. Since animals were not supposed to have souls — at least in the Christian interpretation — any theory requiring an animal ancestry for mankind seemed to deny our spiritual

status.[13] Many giants of nineteenth century science, such as Charles Lyell, refused to accept the possibility of the evolution of species largely because of their "fears for the status of the human soul."[14]

The intention here, however, is not merely to identify conflicts between competing theories, but also to label the failure to comport with evolution as a shortcoming of the soul-theory. Beliefs, whatever their religious significance, "that are flatly incompatible with evolution…are therefore intellectually untenable in spite of their emotional appeal."[15] We want, in other words, to prioritize the respective theories, and expect evolutionism to be the more highly regarded. But in the event of conflict why should it be the burden of the soul-theorist to conform to the context of evolutionary theory and never the other way around? In what way can this be argued to be more than an arbitrary choice made by someone already predisposed to favor evolutionism?

We can begin by repeating the simple observation that if two statements about the same thing are in genuine (and not apparent) conflict, they both cannot be true in the same sense: reality cannot be self-refuting. All things being equal, we should therefore prefer that

---

[13]Peter J. Bowler, *Theories of Human Evolution* (Baltimore, MD: Johns Hopkins University Press, 1986), 1.

[14]Ibid., 2.

[15]Simpson, *Meaning of Evolution*, 5.

theory "whose falsity has not been established."[16] This means that we should never favor a falsified theory over a nonfalsified theory. More commonly, the two competing theories are both presumed or known to be false, leaving us to invest our energies in that one which is less false.

Karl Popper holds that while we have no rational basis upon which to conclude that a statement is true, "we can have strong and reasonably good arguments for claiming that we may have made progress towards the truth."[17] Thus the "search for verisimilitude is a clearer and a more realistic aim than the search for truth."[18]

> Intuitively speaking, a theory T_1 has less verisimilitude than a theory T_2 if and only if (a) their truth contents and falsity contents (or their measures) are comparable, and either (b) the truth content, but not the falsity content, of T_1 is smaller than that of T_2, or else (c) the truth content of T_1 is not greater than that of T_2, but its falsity content is greater. In brief, we say that T_2 is nearer to the truth, or more similar to the truth, than T_1, if and only if more true statements follow from it, but not more false statements, or at least equally many true statements but fewer false statements.[19]

Ideally the preferred theory is one which both increases truth content[20] and decreases falsity content.[21]

---

[16]Popper, *Objective Knowledge*, 8.

[17]Ibid., 57-58.

[18]Ibid., 57.

[19]Ibid., 52.

[20]Truth content is defined as "The class of all the *true* statements which follow from a given statement (or which belong to a given deductive

The procedure to minimize the falsity content is to subject the statement to "severe tests." A statement has been severely tested when its more unlikely predictions have been challenged. A "well corroborated" statement has successfully survived a series of severe tests. Popperian corroboration is a conclusion based upon a past history of testing an hypothesis.[22]

It is difficult to submit the basic evolutionary premise to a severe test given the time scales involved and the retrospective quality of most of its actions. As such, it is open to all the weaknesses typical of an historical thesis (e.g., how do you "severely test" a thesis about exactly why Caesar crossed the Rubicon?). Nevertheless, relative to its

---

system) and which are not tautological." Id. at 48. Tichý formalizes this definition as: The *truth content* A_T *of A* is $Cn(A) \cap T$. Pavel Tichý, "On Popper's Definitions of Verisimilitude," *British Journal for the Philosophy of Science*, 25(1974), 156.

[21]Falsity content is defined as "The class of false statements entailed by a statement — the subclass of its content which consists of exactly all those statements which are false." Popper, *Objective Knowledge*, 48. Tichý ("On Popper's Definitions," 156) formalizes this definition as: The *falsity content* A_F *of A* is the relative content of *A* given A_T, i.e., A, A_T, where the *relative content of A, B of A given B* is $Cn(A \cup B) - Cn(B)$.

[22]Popper, *Objective Knowledge*, 18: "Corroboration (or degree of corroboration) is thus an evaluating *report of past performance*.... Being a report of past performance only, it has to do with a situation which may lead to preferring some theories to others. *But it says nothing whatever about future performance, or about the 'reliability' of a theory.*"

theoretical competitors accounting for the biological world, the basic evolutionary premise is well-corroborated.[23]

Further, we can assert that the thesis of evolution is, despite these philosophical limitations, *better* corroborated than are defenses of the soul.[24] This in itself does not mean that these defenses are less true (or more false). Theories about the soul are even more rarely severely tested and hence are poorly corroborated even if they should happen to be true. But the burden here is to contrast two theories, and it would be irrational, without other justifications, to fail to use the better corroborated theory as the standard. This step entails a conversion of a fact about corroboration into a conclusion about verisimilitude: the greater the corroboration, the greater the verisimilitude. This

---

[23]Ian Tattersall, *The Fossil Trail* (New York: Oxford University Press, 1995).

[24]"Evolution is a fact.... [The] notion that life evolved is about as well established as any fact of science." Salvador E. Luria, Stephen Jay Gould, and Sam Singer, *A View of Life* (Menlo Park, CA: Benjamin/Cummings, 1981), 574. These authors argue that "Evolution is proved by its imperfections" (p. 581). They are here turning the creationists' primary argument from design against them:

> For creationists, adaptation reflects the wisdom of God and the harmony of His world. Exquisite adaptation is the closest thing to perfection that organisms display, and perfection need not have a history.... Evolution, however, is the theory that life has a *history* of continuous change through time.... History is reflected in imperfection, in those aspects of organic form and distribution that do *not* make sense under any notion of a perfect or optimal world" (pp. 576-577).

relationship is not unproblematic, but it is one explicitly endorsed by Popper himself;

> As I have often emphasized, a statement like 'The theory *a* is nearer to the truth than the competing theory *b*' is never demonstrable, but may be nevertheless asserted as a conjecture, strongly arguable for or against on the basis of (1) a comparison of the logical strength of the two theories and (2) a comparison of the state of their critical discussion, including the severity of tests which they have passed or failed. ((2) can also be described as a comparison of their degree of corroboration.)[25]

Given the uncontested greater corroboration of evolutionary theory, Popper finds it not unreasonable to use that outcome to conclude that that theory possesses greater verisimilitude relative to soul-theories. Consequently, where an irresolvable conflict arises between the two the field should presumptively be yielded to the more verisimilitudinous, that is, to evolutionism. The intellectually proper response for the soul-theorists in face of such conflict is to revise his own model rather than to attack the evolutionary competitor.

---

[25]Karl Popper, "A Note on Verisimilitude," *British Journal for the Philosophy of Science*, 27(1976), 158. Cf. Stephen Thornton, "Karl Popper: Probability, Knowledge and Verisimilitude," *Stanford Encyclopedia of Philosophy Stanford Encyclopedia of Philosophy* (Spring 2000 Edition), Edward N. Zalta (ed.), URL = http://plato.stanford.edu/archives/spring2000/entries/popper: "Scientific progress, in other words, could now be represented as progress *towards* the truth, and experimental corroboration could be seen an *indicator* of verisimilitude."

For these reasons we will ensconce the better corroborated evolutionary gradualism as the standard with which a rational defense of the self-preserving soul must comport.

2. Potential weaknesses of this method

The enterprise to ensconce evolutionary theory as the "gold standard" by means of Popper's philosophy of science is not without weaknesses. These include the issues of potential relativism, theory type, induction, and Miller's refutation.

(a). Theoretical relativism

Advocates of the soul have often tried to turn a Popperian perspective against itself. If no statement can be known to be true, then any statement which is not known to be false should be acceptable. The conscientious philosopher must admit that many soul-theories have not been shown to be false, but only that they are weaker theories than their competitors.[26] So perhaps creationist and evolutionary theories, since they are both unfalsified, should both be taught in the schools.[27] But Popper has considered and rejected this line of thinking:

---

[26]Derek Parfit, for example, allows that the existence of "Cartesian Pure Ego, or spiritual substance" is not a "senseless" belief, but it is one for which convincing evidence is missing, and as such is a theory which should be dismissed in favor of more justified competitors. Parfit, R*easons and Persons*, 224.

[27]Stanley cogently outlines how this end is achieved. In essence, while the evolutionist is expected to "prove" his case, the creationist need only create doubt in the mind of the audience. He then capitalizes on

> *From the point of view of objective knowledge, all theories therefore remain conjectural.* From the point of view of practical life, they may be far better discussed, criticized, and tested than anything we are accustomed to act upon, and to regard as certain.
>
> There is no clash between the thesis that all objective knowledge is objectively conjectural, and the fact that we accept much of it not merely as 'practically certain', but as certain in an extraordinarily highly qualified sense; that is, as much better tested than many theories which we constantly trust our lives to (such as that the floor will not collapse, or that we will not be bitten by a poisonous snake).[28]

The fact that, by Popperian lights, neither soul-theories nor evolutionary models are "true" is thus no ground to dismiss a potentially invidious comparison, for they are not being compared on truth but on verisimilitude.

(b). Theory type

Popper's method of theory selection in the present context is sustainable only if the theories being compared are of the same type. That is to say, that both make empirical assertions which can be loosely termed "scientific." Although this condition is necessarily met by all theories of organic evolution, the same is not true of all soul-theories. Our attention is restricted to those soul-theories which have invited

---

this uncertainty and the innate sense of fairness "because the question he can then ask is whether the audience does not find such debates useful. Should not both sides of the question be widely aired and debated?" Stanley, *The New Evolutionary Timetable*, 177.

[28]Popper, *Objective Knowledge*, 80.

treatment as rational arguments with empirical implications. They are intended to be "scientific conjectures," to be tested in some sense of the word, just like evolutionary theories. Even purely metaphysical theories can be tested at least in the sense of their conformity to rules of logical analysis.

In any event we should recognize that some imaginable soul-theories will fail to take the form required for the comparison proposed here.[29] Our method does not embrace these, and hence neither will the conclusions.

(c). Miller's refutation of verisimilitude

Popper's articulation of the concept of verisimilitude ignited an ongoing critical response. We can read this literature as an effort to answer two questions: First, is verisimilitude itself, as a philosophical concept, defensible? And if so, is Popper's version of verisimilitude, the one adopted here, viable?

The theory of verisimilitude is admittedly unsettled. The concept seems to capture an attractive intuition, yet attempts to formalize the details have fallen short. The cornerstone of the critical literature was

---

[29]As an example of a soul-theory which is not amenable to the present project, I would point to the Bible. To the extent that it contains statements about the soul *per se*, these are asserted without defense. While it is possible to evaluate the success of commentators' extrapolations from Biblical verses (e.g., Augustine, Aquinas), the rendered conclusion applies to that interpreter, and not the Bible.

written by David Miller.[30] Using logic statements I will not pretend to be able to follow, Miller concludes that "no two distinct axiomatisable theories may be compared for verisimilitude unless they are both true. But in this case we shall hardly need the sophistications of verisimilitude to tell us which of the systems we would in principle prefer."[31] "The only good fit is a perfect fit."[32] The first blow to verisimilitude, then, is Miller's claim that, as defined by Popper, two false theories cannot be ranked according to their differing proximations to the truth.

The second identified weakness was the vexing argument that "Any verisimilar ordering of theories…has to be bound to the language of those theories…. [Verisimilitude] is not invariant under translation into equivalent theories in a different language."[33] That a theory's verisimilitude should be language dependent is initially disturbing to anyone who regards "Truth" with a capital-T. If there is one Truth, how

---

[30]David Miller, "Popper's Qualitative Theory of Verisimilitude," *British Journal for the Philosophy of Science*, 25(1974), 166-177. Miller's paper is often considered in conjunction with in conjunction with another simultaneously published by Pavel Tichý, "On Popper's Definitions of Verisimilitude."

[31]Ibid., 174.

[32]David Miller, quoted by Chris Brink, "Verisimilitude: Views and Reviews," *History and Philosophy of Logic*, 10(1989), 186.

[33]Brink, "Verisimilitude: Views and Reviews," 186.

can a theory's approach to that Truth depend upon the language in which the theory is stated?

Although Brink writes that with these arguments "Miller… questions the worth of the whole verisimilitude enterprise,"[34] it is not immediately obvious that Miller intended to discredit the intention of verisimilitude, or only to critique Popper's particular version thereof. Miller later takes a stab of his own at formalizing the concept, so this larger goal either was not his intention or one which he later disavowed. Still, according to Brink, Miller's "has become the canonical objection against the whole verisimilitude enterprise."[35]

In response to Miller, Popper denied that the linguistic "'relativisation' of theories" was a difficulty unanticipated by his own work.

> Thus Miller's interesting results (assumed here to be correct) interfere neither with the idea of an approach to truth nor with my methodology…. But his results may possibly force us to an historical relativisation of the idea of verisimilitude, in the sense that two historically isolated and different chains of problems *cum* solutions may become comparable with respect to verisimilitude only after the two chains have merged.[36]

---

[34]Ibid.

[35]Ibid.

[36]Popper, "A Note on Verisimilitude," 157.

As I understand the argument, verisimilitude is correctly clarified as not being a property or value possessed by a theory. One cannot say, "Theory X has a lot of verisimilitude." Verisimilitude is a comparative assessment, and hence is always relative to the specific contrasting theory. Soul-theories may suffer in verisimilitude judgments against evolutionary theory, but against some other theory they may do better. As the full theoretical context will determine the proper language of the theories, Miller's observation that linguistic variables impact verisimilitude assessments does not render such comparisons meaningless. One can easily imagine that a theory that has low original success profits immeasurably when rephrased in some other terms.[37] Linguistic relativism, then, constrains the kinds of verisimilitude which may be formally asserted, but is not fatal to the verisimilitude project itself.

Some philosophers (such as Graham Oddie) base their work upon the assertion that there exist "some core examples of absolutely uncontroversial judgements of verisimilitude [and that these] are the

---

[37]As one possible example, the theory of telepathy as an exchange of "mental energies" modeled on radio transmissions has been resoundingly discredited. However, a rephrasing of the theory which adheres closely to the original phenomena and which recasts telepathy as a manifestation of the social psychological phenomena of charisma and telepathy may fare better. Cf. James M. Donovan, "Reinterpreting Telepathy as Unusual Experiences of Empathy and Charisma," *Perceptual and Motor Skills*, 87(1998), 131-146.

yardstick by which conflicting accounts of verisimilitude are to be judged. Most of us would agree that "Given that there are nine planets…the theory that there are eight planets is better than the theory that there are seven."[38] Even granting this intuition, Brink concludes that "it seems unlikely that one single model of verisimilitude will suffice for all contexts."[39] If in most contexts an eight-planet theory is closer to the truth than a seven-planet theory, against other backgrounds the latter might be the better theory. The point to be drawn, however, is not that the lack of a single standard of verisimilitude demonstrates the emptiness of the concept, but only that the verisimilitude of a given theory is relative to our larger intellectual needs at that moment.

Brink characterizes the aim of the authors he discusses as being "to establish…the *respectability* of the verisimilitude enterprise."[40] This much at least we may take as achieved, even if the details of how verisimilitude is to be ascertained. But this is enough for our purposes. The present work need assume nothing more than the intuitive sense of verisimilitude that Popper identifies, and the expectation that verisimilitude is a function of corroboration. Nothing hinges on any particular attempt to formalize the concept.

---

[38]Brink, "Verisimilitude: Views and Reviews," 192-193.

[39]Ibid., 199.

[40]Ibid., 198.

In keeping with Popper's concession Brink concludes his review with the dual observations that "the *respectability* of the verisimilitude" has been largely rehabilitated, but that "it seems unlikely that one single model of verisimilitude will suffice for all contexts."[41] Popper would seemingly have had no problem with this. As we saw above, he conceded that assessments of relative verisimilitudes was not demonstrable, but only a conjecture. The "chief value of the concept is heuristic and intuitive, in which [case] the absence of an adequate formal definition is not an insuperable impediment to its utilisation in the actual appraisal of theories relativised to problems in which we have an interest."[42]

So while it can be perilous to attempt to discuss verisimilitude in the abstract, as though it referred to an invariant absolute, it is much less dubious to invoke the concept in a particular, defined context such as the present thesis. It remains to be seen whether the version of verisimilitude applied here is similarly acceptable.

If verisimilitude is a legitimate philosophical concept, we can turn our review to the second question of whether Popper's version is a proper touchstone for our analysis. We consider here the logical, formal approach to the problem.

---

[41]Brink, "Verisimilitude: Views and Reviews," 198-199.

[42]Thornton, "Karl Popper."

Brink discusses a paper by Schurz and Weingartner which proposes to salvage Popper's idea from the Tichý -Miller inadequacy proof.

> Recall Popper's definition: it says in effect that theory α has greater verisimilitude than theory β iff α has more true consequences and fewer false consequences than β. The Tichý -Miller inadequacy proof capitalises precisely on this aspect of the definition, that the *consequences* of a theory determine its verisimilitude. Here 'consequence' means 'logical consequence', and the central point of Schurz and Weingartner's argument is that a scientist may not *want* all logical consequences: 'The concept of logical consequence in classical logic', they say, 'is broader and more tolerant than that which is actually used by scientists in everyday life.' Some deductive consequences are, to a scientist, simply *irrelevant*, others simply *superfluous*.... Careful analysis of the Tichý -Miller inadequacy proof reveals that it hinges precisely on adding either an irrelevant or a superfluous consequence. Therefore, if we disallow such consequences, the inadequacy proof is blocked and (a version of) Popper's original definition is thereby rehabilitated.[43]

At the very least the invocation of the concept of verisimilitude is not quite the resuscitation of a proven failure as some might assume.

As stated above, however, whatever the fate of the formal casting of verisimilitude, Popper allows that these are intuitive judgments and not deductive conclusions. The central claim which we are adopting from Popper, and which has thus far been

---

[43]Ibid., 196. The paper by Schurz and Weingartner discussed here appears in T.A.F. Kuipers (ed.), *What is Closer-to–the-Truth?* (Amsterdam: Rodopi, 1987).

unaddressed, is the use of corroboration to make assessments of verisimilitude. We turn now to this problem.

(e). Humean inductive skepticism

We have seen that verisimilitude is a credible philosophical concept, and that Miller's critique of Popper's formulation may not be completely fatal to that particular version. But the greatest enduring weakness of Popper's version may be the fact which has been repeatedly mentioned, and upon which the present project particularly depends. Specifically, if the use of corroboration to infer verisimilitude is illegitimate, then the argument to establish evolutionism as the theoretical "gold standard" in the present project falters.

As stated, corroboration is "a conclusion based upon a past history of testing an hypothesis." We saw that Popper regards corroboration as a basis upon which to determine verisimilitude. One implication of a claim that theory *a* has greater verisimilitude than theory *b* is that theory *a* is expected to perform better than *b* on future tests. In other words, corroboration, a summary of past performance, is used to justify judgments about verisimilitude, which carries implications for future performance. This use of past events to predict future outcomes is the gist of inductive reasoning.

Keeping in mind that Popper approved the use of corroboration to infer verisimilitude, Popper also explicitly disallows the kind of inductive

reasoning required of that step. Popper is necessarily wrong on one of these mutually exclusive claims: Either corroboration can be employed, via inductive reasoning, to infer verisimilitude, or, because inductive reasoning is unjustified, the step from corroboration to verisimilitude cannot be taken. If we cannot find good reason to prefer the first over the second, the proposed theory preference method is seriously impacted.

Skepticism about the justifiability of induction originates with David Hume. Hume's "big idea" was that all our knowledge must originate in our sense perceptions. As such, only those ideas which reflect sensations are valid. The idea of "cause and effect" has no such primary sensory correlates. What we *do* have is the perception of constant correlation, which through habit and custom we construe as causes and effects. But this is a psychological convenience, and not a rational conclusion. Induction is just this kind of trying to predict future performance from past observations, and is similarly unjustified.

"Popper is unusual amongst contemporary philosophers in that he *accepts* the validity of the Humean critique of induction, and indeed, goes beyond it in arguing that induction is never actually used by the scientist."[44] Popper believed that one of the virtues of his philosophy was

---

[44]Thornton, "Karl Popper." While it is incontestably true that Popper sides with Hume, it has not been my impression that this stance was "unusual amongst contemporary philosophers."

that it circumvented this logical boundary by building his philosophy upon deduction, not induction. To use corroboration in the way described directly contravenes his expressed intentions. But is induction that much of a problem?

I am insufficiently quelled by force of Hume's argument against induction for two reasons.[45] First, he directs the force of his argument against the use of induction to derive cause and effect relationships. But this is not the only use for induction, and an effective argument against the one cannot be construed, given the latter, a general critique of induction.

I am here specifically thinking of the use of induction, as Aristotle points out, to generate categorical universals.[46] The mere fact that we recognize a sense impression as a "cat," whether or not we are able to use "cat" as a predicate in a cause and effect chain, is attributable only to induction. The stock Humean position is that where we think we see

---

[45]Although Popper and other philosophers find Hume's arguments determinative, others are less certain. Cf. Wesley C. Salmon, "Should We Attempt to Justify Induction," *Philosophical Studies*, 8(1957), 33-48. Salmon concludes that arguments favoring the abandonment of induction are inconclusive, and that the search should continue, given the high price to be paid in terms of conventionalism (relativism).

[46]Aristotle, "Posterior Analytics," translated by G.R.G. Mure, in *The Basic Works of Aristotle*, edited by Richard McKeon (New York: Random House, 1941), 81b5-9.

powers and laws, we will find only habit and custom. Is our recognition of the category "cat" simply this, habit and custom?

Yes and no. There is some variability in category construction, but frankly not so much that that project is best explained by such arbitrary mechanisms.[47] George Lakoff emphasizes that categorization is a creative process, but also that it is neither random nor accidental. If it were, cross-linguistic comprehension would be monstrously difficult if not impossible. Identifiable principles are operable which "motivate" certain outcomes, without determining them.

So perhaps induction does have its uses, even if one of them is not to justify claims of cause and effect. This fact is sufficient to render any universal dismissal of induction untenable. But I am not prepared to immediately concede that there is even this limitation. I certainly cannot hope to dispel this enduring philosophical problem here. But I should indicate the directions I would take were I to approach this task.

I suspect Hume has unreasonably differentiated between time and space, and between units and sets. The problem with recognizing the relationship between A and B should be the same regardless of whether

---

[47]Cf. George Lakoff, *Women, Fire, and Dangerous Things: What Categories Reveal about the Mind* (Chicago: University of Chicago Press, 1987).

they are separated by time or by space.[48] If my arm is hidden by a screen, are you justified in assuming that the arm is there, even though you cannot see it? Perhaps you have seen my arm before; you *know* I have a second arm. So you conclude, although you have no current sensory information to that effect, that I have a second arm.

What if you have never met me before? Can you then rationally conclude that I have a hidden second arm? Not as a deductive certainty. But you can conclude that the existence of such an arm is probable, given the premises that anatomically normal humans have two arms, and that I belong to that class of humans.[49] In that instance it would be rational for you to conclude that I have a second arm, provided it has not been removed. Your expectation may be incorrect, but it is not irrational to hold it. Hume's critique is directed not toward achieving certain knowledge but toward identifying what it is rational to believe.

---

[48]This position is not uncommon among philosophers. Derek Parfit, for example, argues that it is a mistake to distinguish between what is spatially far away and what is temporally remote; neither spatial nor temporal proximity alone justifies disparate treatment of individuals. Parfit, *Reasons and Persons*. Nor is this position alien to physics. The terms are often conjuncted, so that we speak of "the space/time continuum," or use the adjective "spaciotemporal."

[49]The gist of my argument is that by conceptually lumping the phenomena differently than does Hume, the problem he identifies dissolves. A similar attempt to resolve a Humean problem by altering the language used to describe the problem can be found in Terence Penelhum, "Hume on Personal Identity," *Philosophical Review*, 64, 1955, 571-589.

Now, how is a physical screen philosophically different from a temporal screen? In at least this instance, not at all. And it is the latter kind of screen which concerns Hume. Given A now, can you predict B latter? This is analogous to: Given my visible body, can you predict that I have a second arm behind the screen?

If the analogy holds, then they can be solved similarly.[50] How did we solve the first? By appeal to the *set* which constitutes the body. The arm behind the screen is not an independent thing; rather, it is part of the set of things which together comprise the "human body." To expect the arm behind the screen is not to infer an independent thing from some other independent thing, leaving us with the Humean problem of how to relate the two. Rather, the arm is a member of the set. The visible body invokes the entire set. Our familiarity with the set allows us to predict or expect what we will find behind the screen.

So too by things separated temporally. Hume casts A and B as being independent events which are problematically related. And we can see why he might think that: our senses perceive A and B as discrete impressions. But so too are the sense impressions of the body and hidden arm separate events; would he likewise argue that we have no

---

[50]This conceptual linkage of space and time has Humean precedence. Hume gives the same kind of account of both, leading us to think that what is true of one must, broadly considered, be true of the other.

basis to claim that body-with-hidden-arm and the body-now-with-two-arms- revealed are the same?

A and B should be understood as a set. A is only one part of the set; B is the other, obscured by a temporal curtain. But the set is invoked in its entirety, even by only a part of it. We can then infer B from A in the same way as we can infer the second arm from the visible body. When the sciences are investigating why A is the "cause" of B, philosophically this would be to explain why A and B are a set.

That, at least, is where I would begin to dissect Hume's problem of induction. I am sure I have been preceded in this direction by others presently unknown to me. But because I do not view the problem of induction as fatal, I am not hesitant to conclude that it is not obviously irrational to use corroboration inductively to infer verisimilitude. Others will disagree.

In summary, we intend to use Popperian verisimilitude and corroboration to prioritize theories. The practical outcome is that evolutionary theories are always superior to soul-theories because the former are better corroborated, and thereby presumptively possess greater verisimilitude. This method is defensible if soul-theories are of the same type as evolutionary theories, facilitating comparison; if verisimilitude is not demonstrably an illegitimate philosophical concept; and if the use of induction to use corroboration to draw conclusions

about future performance is not immediately unwarranted. Arguments have been offered to show that these three conditions apply. Consequently, evolutionary theory shall function herein as the standard to which soul-theories must conform. Where these two theories are in conflict, it is the intellectual burden of the soul-theorist to revise his own position.

C. Possible Arguments to Avoid the Conflict of Theories

We have seen that the rational defense of the self-preserving soul requires a confrontation with evolutionary theory, and that this defense is strategically disadvantaged in this intellectual contest. As a result, many soul-theorists will wish to evade the responsibility of reflecting the implications of evolutionary biology. This section rebuts three possible evasive arguments.

1. <u>Religious stances are exempt from empirical criticism</u>

The first strategy is to exempt religious claims from any responsibility to comport with even the most secure of scientific knowledge. This resort to a religious gloss to avoid conflict is relevant because although the present piece intends to be a philosophical work, the "soul" is often appropriated as an object of religious significance. To the extent that "soul" is thus an object of *belief*, the objection might merit deep consideration: Is it irrational to hold religious beliefs which contradict other areas of our knowledge?

But that is not the problem here. We do not set out to determine what propositions should be believed, but only to ascertain which propositions it is rational to believe. There may be other grounds to believe a claim than its rationality, but we do not investigate the possibility. Yet to the extent that the belief in the "soul" is *rationally defended* and not merely asserted as elemental dogma, the argument becomes philosophical if not scientific and therefore open to the same critical scrutiny as any similar proposition.

2. "Humans" versus "*Homo sapiens*"

A second possible argument to relieve soul-theorists from the duty to incorporate the implications of evolutionary thinking into their models is based upon the failure to meet the condition required of any comparison, namely that the two statements are "about the same thing."

Strictly speaking, biology speaks of *Homo sapiens*[51] while soul-theorists refer to the "human being." If the terms are not synonyms, or at least sufficiently similar for purposes of a critical comparison, then the requirement that the statements not conflict is removed.

---

[51]Although it can be cavalier to generalize about these matters, in general the genus *Homo* is differentiated by its average brain size in excess of 800 cc., a more reduced face and teeth with a larger body size relative to the other geni within Hominidae. Within *Homo*, *Homo sapiens* has an even larger brain size, more than 1000 cc., with further reduction in teeth and face. Luria, Gould, and Singer, *A View of Life*, 665.

The problem of demarcation between animal and human is philosophically vague, and hence scientifically inspecific. Some workers might think that they can avoid this problem by comparing only clear exemplars of the invoked categories. A dog, for instance, is surely a clear example of an "animal" by any standard, while a philosopher is a similarly uncontested instance of "human." Restricted to these terms, there is in fact little problem, so long as the conclusions are couched in the same terms, i.e., of "dog soul" and "philosopher soul." But if we wish to generalize to the broader categories of animal and human, the sorting problem re-presents itself.

The "term 'human' had been deprived of any hope of precise definition as soon as it was realized that people have an evolutionary history."[52] One could thereafter only argue that the distinction between animal and human was unproblematic if one was also prepared to assert that humans are not animals. That claim requires minimally the proposition that humans are the outcome of a separate creation event and share no genealogical ties to other lifeforms.

The anatomical similarities are far too great to support such a distinctive origin for humans. Most educated persons begin with the assumption that humans are part of the evolutionary process, assigning

---

[52]Tattersall, *The Fossil Trail*, 114. This problem might be analogized with the difficulty in pinpointing the precise instant when a light slowly cycling through the visible spectrum suddenly became "red."

them membership in the Animal Kingdom. Whatever they may be in addition to this, humans have their origin here, and are thereby animals.[53] To contrast animals with humans, then, is perhaps disingenuous from the outset.

But the distinction is marked; our task is to find out how. The first attempted clarification offered might be that "animal" means "non-*Homo sapiens*." Those organisms belonging to the species *Homo sapiens* are therefore "humans." There is a strong sense in which we take this alignment to be accurate. We have two reasons though to reject the contention that "human = *Homo sapiens*."

First, we should, if "human" and "*Homo sapiens*" are synonymous, be able to manipulate them identically. That is, we should expect the two to be linguistically interchangeable. But this is not the case. Consistently we find that the criteria for "human" are behavioral, while those for "*Homo*" are biological.[54]

---

[53]There is some reason to think that the animality of humans was less distasteful than the assumption that it was the apes particularly to which we were related. Some earlier evolutionary scenarios went to great extremes to disavow that particular ancestry while doing nothing to mitigate the conclusion that, at base, we are animals. Cf. Bowler, *Theories*, Chap. 5.

[54]For example, Harry Stack Sullivan writes that

> the thing which distinguishes the human being from the human animal is the incorporation in the poor human animal of vast amounts of culture, of socially meaningful, rather than biologically meaningful entities, which exert very

If aliens who landed on our planet looked very much like us, we would call them "humanoid," not "human." They might earn that label, though, by stopping to pet the kitten and being moved by the plight of starving Africans. Nothing they could do, however, could earn them the label "*Homo*," even metaphorically. Contrarily, someone is called "inhuman" based solely upon behavior: the aliens who vivisect their conscious abductees, for example.

This same point can be illustrated by considering what we mean when we discuss "human nature."

> **Human nature**, in general terms, denotes the nature of man, with more especial reference to his personality and/or character *as acquired in the course of socialization* and often with further reference to aspects of human potential and powers of development.[55]

This definition of "human nature" is typical in that that which is "human" emerges from the organism's socialization into a culture. "Philosophies of human nature reflect beliefs about what people are like *after* they have moved through a lengthy socialization process. The concept does *not* attempt to reflect beliefs about inherent or innate

---

> powerful influence on all subsequent performances of the creature.

Harry Stack Sullivan, "The Illusion of Personal Individuality," *Psychiatry*, 13(1950), 323.

[55]E.H. Volkart, "Human Nature," in *A Dictionary of the Social Sciences*, edited by J. Gould and W.L. Kolb (New York, Free Press, 1964), 306 (emphasis added).

qualities."[56] Work in this area does recognize constraints that are *biological* rather than *social*, but these are not collected under the category of "human nature" but instead within "psychic unity."

This perspective holds that the organism left to develop in isolation might survive but would never be quite human, regardless of its genotype.[57] That is, *Homo* is not necessarily human where it lacks enculturation. Alternatively, again regardless of the genotype, an enculturated organism may be definitionally human. If culture and being human are strictly correlated, "then some chimpanzees may be human or human-like."[58]

If to be or not to be human (and thus, in the latter instance, to be animal, whether *Homo* or something else) is contingent upon a process of socialization, then judgments about a creature's humanness are not

---

[56]L.S. Wrightsman, *Assumptions about Human Nature*, 2nd ed. (Newbury Park: Sage, 1992), 46.

[57]An early articulation of this position is to be found in Aristotle's *Politics*:

> an individual *incapable* of membership of a *polis* is not, strictly speaking, a human being, but rather a (non-human) animal, while one who is self-sufficient apart from the *polis* is superhuman, or, as Aristotle puts it, a god.... [One] cannot be a human being except in the context of a *polis*.

C.C.W. Taylor, "Politics," in *The Cambridge Companion to Aristotle*, edited by Jonathan Barnes (New York: Cambridge University Press, 1995), 239.

[58]William Noble and Iain Davidson, *Human Evolution, Language and Mind* (Cambridge: Cambridge University Press, 1996), 83.

equivalent to a decision about the creature's species but begin to draw in other considerations like behavior, which is to say that "human ≠ *Homo sapiens*."

Judgments about whether a particular token of *Homo* qualifies as human can be very arbitrary and historically variable. Prior to the Middle Ages humans and animals were viewed as biologically permeable categories, meaning that individuals could just as easily move in one direction as the other: Humans become animals (werewolves, vampires), animals become humans (the characters of Aesop's fables). Later, however, behavioral category markers opened the possibility that even certain members of *Homo* routinely failed to qualify as human. If being "rational," for example, is the definitional human trait, then persons who act in ways judged "irrational" (e.g., women in general, peasants in French folk tales) may come off as less than human.[59] This trend can so overexercise itself that by tighter and tighter constriction no actual humans exist: If true rationality is the possession of only the theoretical philosopher, only he might be human. But in the last analysis true

---

[59]Joyce E. Salisbury, "Human Beasts and Bestial Humans in the Middle Ages," In *Animal Acts: Configuring the Human in Western History*, edited by Jennifer Ham and Matthew Senior (New York: Routledge, 1997), 15.

knowledge is an impossible goal that no one achieves, so no one at all is human.[60]

No ability deemed "human" is possessed by every living representative of *Homo* because, as we saw, humanness is an achievement of culture, not an inheritance from biology. Consequently might some of our friends and relatives really be "animals," having failed to be properly socialized? And what, it still remains to be determined, constitutes "proper" socialization?

We should concede that "human" is not synonymous with *Homo*. This fact does not necessarily render invalid the equation "human = *Homo sapiens*." Perhaps the two terms mark different aspects of the same thing, the way "cylinder closed at one end" and "holder of liquids" refer to different aspects of a cup. The obstacle here is that whatever criteria used to characterize one will drastically impact the other.

Start with the fact that whatever behavioral trait we choose as essential to the concept of "human," we will end by withholding this label from many of our fellow *Homo sapiens*. We have already seen this process at work if we require of humans that they be rational. In another example Stephen Mithen characterizes a distinguishing feature

---

[60]Gregory B. Stone, "The Philosophical Beast: On Boccaccio's Tale of Cimone," in *Animal Acts: Configuring the Human in Western History*, edited by Jennifer Ham and Matthew Senior (New York: Routledge, 1997).

of the human mind as one that possesses the ability of cognitive fluidity.[61] That is, the separate modules of cognition can have their results accessed by other modules, so that intelligence is general and not specific. His main concern is to differentiate *Homo sapiens* from earlier representatives of its evolutionary lineage. But his criterion has the undiscussed implication that living *sapiens* lacking this ability (e.g., "idiot savants," a form of autism) must also not qualify as being fully human. The point is that whatever criterion one asserts to characterize "human" over and above animal, some members of *Homo* will not possess that quality or trait. If standards are strictly applied these entities are not human, but instead animal despite their membership in the "home species."

This argument does not imply that the human/animal differentiation cannot be made in principle. Rather, it suggests that a principled divisioning will be messier than most people assume. The strong temptation will be to resort to special pleading to "correct" any result contrary to that generated by pretheoretical expectations (e.g., how

---

[61]Steven Mithen, *The Prehistory of the Mind: The Cognitive Origins of Art, Religion and Science* (London: Thames and Hudson, 1996).

to define "human" so that it fully captures our intuitions but which does not systematically exclude, say, the mentally ill or the fetus?[62]).

So the brute fact of biological variability means that whatever behavioral criterion we select for "human," some specimens of *Homo* will lack this trait (and, depending on the trait, some non-*Homo* might possess it: our aliens, for example). Interestingly, the same applies even if we restrict ourselves to the biological criteria implied by the term *Homo*. What, *exactly*, earns someone that label? The number of chromosomes? Some people have extras: are they therefore not human?

Enough information has been presented to show that the commonly presumed equation "human = *Homo sapiens*" is inadequate. "Human" refers to behaviors; *Homo* to biology. Many *Homo* are clearly not human, and some humans are contestably non-*Homo.* The issue then becomes whether the terms are so divergent that statements about the one have no rational duty to mesh with those of the other.

An adequate response to this problem depends upon the details of the specific soul model being defended. For instance, the failure of the equation is a nonissue for those models who fail to make this distinction, ascribing equivalent souls to both categories of animal and human. Such models, for example, allow for transmigration of souls between

---

[62]Thomson, *Realm of Rights*, 288-293, outlines the problems associated with arguing that the fetus is "human" and therefore possessed of rights.

humans and animals. These models are probably a minority, and a different reaction is required for those which do make the distinction between human and animal.

Viewed differently, the failure of the equation could remove the imperative for the soul model to precisely track *anatomical* evolution. Since "people began to behave in a modern fashion only well after they had assumed modern anatomy,"[63] if the souled difference attaches to modern behavior then "human" postdates attainment of *Homo* status. To claim this potential refuge, however, the soul-theorist must identify the behavior whose origin it wishes to assign to the soul, and argue that that behavior is both modern and unique.

The soul-theorist must identify what difference the soul is supposed to be making. That is, the soulless *Homo* differs from the souled human in precisely what attributes? A difference must exist lest the soul be an empty ascription. If the difference is not correlated with anatomy, where should we look?

Many soul-theories would probably point to some member of a suite of psychological and cognitive features such as emotions or moral sense. "The opponents of Darwinism stressed the unique nature of our mental and moral faculties to defend the traditional belief that they could

---

[63]Tattersall, *The Fossil Trail*, 245.

only have arisen by divine creation."[64] Indeed, Alfred Russell Wallace, credited as the co-discoverer with Darwin of natural selection, "maintained that a number of our mental and physical characters could have been formed only by direct supernatural intervention in human evolution."[65] The obvious response is that these things, too, have probably evolved. Darwin's own opinion was that the first foundation or origin of the moral sense lies in the social instincts, including sympathy; and these instincts no doubt were primarily gained, as in the case of the lower animals, through natural selection.[66] Social cooperation,[67]

---

[64]Bowler, *Theories*, 150.

[65]Ibid., 159. Wallace's stance should not be construed as groundless supernaturalism.

> Hence, Wallace's dilemma: all "savages," from our actual ancestors to modern survivors, had brains fully capable of developing and appreciating all the finest subtleties of European art, morality and philosophy, yet they used, in the state of nature, only the tiniest fraction of that capacity in constructing their rudimentary cultures, with impoverished languages and repugnant morality.
>
> But natural selection can only fashion a feature for immediate use. The brain is vastly overdesigned for what it accomplished in primitive society; thus, natural selection could not have built it.

Gould, *The Panda's Thumb*, 55.

[66]Charles Darwin, *The Descent of Man, and Selection in Relation to Sex* (Princeton, NJ: Princeton University Press, 1981), Vol. II, 394.

[67]Robert Axelrod, *The Evolution of Cooperation* (New York: Basic Books, 1984).

altruism,[68] and moral systems generally[69] have all yielded to contemporary evolutionary analyses.

The result for these soul models, then, is that while they may be absolved of the responsibility of intermeshing their theories with *Homo*'s evolving anatomy, they are not similarly relieved of the task of accounting for evolutionary incremental development of the alternative trait marking the appearance of the soul.

To avoid this obstacle entirely they would have to characterize the soul in terms of traits which clearly appear in toto, which would still do justice to the idea of the soul. I can think of no such trait, but that does not mean that someone else cannot propose one. Failing this advance, the problem for soul-theorists remains the same: how to reconcile incremental evolutionary development with a soul which has no precursors. All that differs here is that it is not anatomical development which provides the point of contrast, but development of some psychological or behavioral feature identified by the soul-theory itself with evolution.

---

[68]Edward O. Wilson, *On Human Nature* (New York: Bantom Books, 1978).

[69]Richard D. Alexander, *The Biology of Moral Systems* (New York: Aldine de Gruyter, 1987). "Darwin argued that any social animal would acquire a moral sense once its intellect had developed to a certain level." Bowler, *Theories*, 157.

This third argument to spare soul-theorists from critique, that their statements are not about the same thing as evolutionary statements, can be successful in refocusing what the evolutionary statements are about (e.g., whether we should be concerned with anatomical or instead with psychological evolutionary development), but fails to avoid the confrontation itself.

3. Argument *ex consensu gentium*

A third argument to avoid the conflict acknowledges that evolutionary theory is legitimate, and that evolutionary and soul-theories are "about the same thing." The tack here seeks to find endorsement of the rationality of the theory (and parasitically, its truth) in its wide popularity. Such a defense takes the form of an invocation of the "argument *ex consensu gentium*." For example, Judith Thomson defends a position by claiming that "the fact that many people think it true surely is some reason to think it true."[70]

In the present context this argument reasons that the object of belief (e.g., the postmortem self-preserving soul), being as it is an object of religious belief, is self-validating. That is, due to the inherently rational nature of human beings they would not so persistently and

---

[70]Thomson, *The Realm of Rights*, 80. Parfit (*Reasons and Persons*, 186), puts this argument in the mouths of imagined opponents, while Simpson (*Meaning of Evolution*, 147, 345) reminds his readers that "repetition does not establish truth," and that, on the contrary, many "things that everyone seems to know [are] not true."

universally believe that which is not true, and therefore the belief in a soul is true because it is believed in this way. If accepted, this argument renders a reconciliation between evolutionary and soul-theories an interesting intellectual problem, but not one which informs the truth of either. Each theory is independently certified in its own right (evolutionary theory by data, soul-theory by acclaim).

Even if one were inclined to accept this characterization of the human being as inherently rational — a position one might regard as untenable after Freud's "excavations" into depth psychology — this argument still fails because it is simply false. Popularity is a proof of neither rationality nor truth.[71] Many popular beliefs have been later shown to be false. False popular beliefs have included the innate biological inferiority of Africans and the flatness of the earth, not to mention "urban legends" like Bubble-Yum being made out of spider webs. The existence of a class of false popular beliefs refutes the claim that popularity confers credibility, either in the form or rationality or truth. It is never the responsibility of reality to conform to our expectations; instead, our beliefs must alter to more truthfully reflect what is real.

---

[71]This same point was defended much earlier by Aristotle: "popular acceptance or rejection is no criterion of a basic truth." "Posterior Analytics," I,6,74b24.

D. Fundamentals of Evolutionary Theory

Part II of this thesis closely examines some candidate soul-theories. This section reviews the principle assertions of evolutionary theory.

1. Evolution, natural selection, and Darwinism

"Perhaps no other concept in the history of science has been so simple in construction — a few evident facts and one obvious conclusion — yet so difficult to transmit for general understanding."[72] This characterization of the theory of evolution does not exaggerate. On the one hand, the argument for evolution is almost skeletal:[73]

1) All organisms within a species vary one from the other;

2) Some of these variations are inherited;

3) All organisms produce more offspring than can survive to become reproducing adults; and

4) If organisms vary and not all can survive, on the average survivors will possess those heritable variations that increase adaptation to local environments.

It is useful at this point to distinguish between three terms that are often used interchangeably: natural selection, evolution, and Darwinism. *Evolution* is merely "historical change in the programmed set

---

[72]Luria, Gould, and Singer, *A View of Life*, 597.

[73]Ibid., 582.

of instructions for building an organism."[74] This entails only the idea that over time the set of inherited traits characterizing a breeding population (e.g., a species) can change. This idea was not controversial by the time Darwin published his massively influential *The Origin of Species* in 1859. The principle achievement of this book in its own day was to establish the credibility of general evolution beyond all reasonable dispute; the more novel facets of the treatise were largely overlooked.

What *was* controversial in Darwin's work was his posited mechanism for evolution. The sketch above is his reasoning for the operation of *Natural Selection*, that is, the "preservation of favourable variations and the rejection of injurious variations."[75] Earlier evolutionary theories (e.g., Lamarck's inheritance of acquired characteristics) had implied that the course of evolution was either directed (by God, perhaps) or involved the unfolding of internal tendencies toward a predetermined goal (orthogenesis). By contrast, natural selection suggested that evolution was (1) purposeless, in that there is no higher principle operating in nature; (2) not inherently progressive, meaning that "there is no ladder of life or chain of being" of

---

[74]Ibid., 574.

[75]Charles Darwin, *The Origin of Species by Means of Natural Selection* (New York: Avenel Books, 1979), 131.

which man is the end product, intended or otherwise; and (3) a philosophical materialism.[76]

This last point is particularly relevant to our present discussion. Darwin advocated a thoroughly naturalistic account of life, thus denying one of the deepest traditions of Western thought: the separation of mind and matter, with higher status for mind. He viewed the human mind as a natural outcome of selective purposes for a large and complex brain. He granted humans no special status and viewed us as inextricably connected with all life. Darwin did not set out to demolish anyone's religious convictions[77]; he merely wished to assert that divine causes and attributes had no place in scientific investigations. He saw the human brain as an object for scientific study, just as any other part of nature.[78] To the extent that "mind" is synonymous with "soul" (a legacy gifted us by Descartes), the revolutionary impact of natural selection becomes obvious.

Just as it is possible to be an evolutionist without accepting natural selection, it is possible to be an evolutionary natural selectionist without being a *Darwinian*. In addition to the basic requirements of

---

[76]Luria, Gould, and Singer, *A View of Life*, 584.

[77]Indeed, by all accounts his wife was a devout Christian who was deeply troubled by her husband's writings.

[78]Luria, Gould, and Singer, *A View of Life*, 585-586.

natural selection, Darwin's formulation also required (1) that variation be random; (2) that substantial amounts of inherited variation be present in breeding populations; and (3) variation occur in small steps.[79] Darwin reiterates the last point in these words:

> As natural selection acts solely by accumulating slight, successive, favourable variations, it can produce no great or sudden modification; it can act only by very short and slow steps. Hence the canon of 'Natura non facit saltum,' which every fresh addition to our knowledge tends to make more strictly correct, is on this theory simply intelligible. We can plainly see why nature is prodigal in variety, though niggard in innovation. But why this should be a law of nature if each species has been independently created, no man can explain.[80]

It is the stipulated requirement of gradualism that distinguishes classic Darwinism from other forms of natural selection evolutionary theories such as punctuated equilibria.

Having distinguished these three terms, we can assess the relative significance of each for the present study. The fact of evolution *per se* is completely uncontroversial and uncontested. The so-called "anti-evolutionists" do not deny that evolution occurs; rather, they deny that this is how species form. They accept "microevolution," defined as "the study of changes in gene frequencies within single populations."[81]

---

[79]Ibid., 583.

[80]Darwin, *Origin of Species*, 444-445.

[81]Luria, Gould, and Singer, *A View of Life*, 601.

Microevolution accounts for cases like the peppered-to-black changes of the moth *Biston betularia* as its natural environment became more polluted and it needed to alter its color to maintain its protective camouflage.[82] Anti-evolutionists reserve their objections for theories of macroevolution,[83] the "large-scale events" which operate on a vast temporal scale effecting qualitative changes (i.e., speciation).[84]

We shall not be using "evolution" in this generic sense. Rather, it will always be restricted to include at least the idea of natural selection. This limitation may seem obvious today. Contemporary readers might be surprised to learn that it was not until 1950 that Darwin's mechanism of natural selection became widely accepted in the form of the "Modern Synthesis."

Traditional conceptions of natural selection build upon random genetic variation by favoring the variants that are more successful in the

---

[82]Thus, for example, creationist Hank Hanegraaff (*The FACE That Demonstrates the Farce of Evolution*, [Nashville, TN: Word Publishing, 1998], 172) suggests that "microevolution," which generates the perfectly acceptable transitions evidenced in dogbreeding, is "perhaps misnamed." His point seems to be that it would be easier to be an "anti-evolutionist" if the term "evolution" did not encompass phenomena both objectionable (like macroevolution) and nonobjectionable (microevolution).

[83]The proposal in Kansas to eliminate evolution from the school curricula and mandatory testing was directed explicitly at claims of macroevolution. William K. Piotrowski, "The Kansas Compromise," *Religion in the News*, 2(3)(1999), 10-12.

[84]Luria, Gould, and Singer, *A View of Life*, 629.

immediate environment. Necessarily the end-product is imperfect, since the inherited qualities are selected as successful in the *parent's* environment, and not the offspring's. The evolutionary process is therefore always one step behind the environment, moving the species along.

The essence of Darwin's theory lies in his claim that natural selection is the directing or creative force in evolution, not merely in the assertion that natural selection operates. Pre-Darwinian creationists were quite comfortable with natural selection as a mechanism for the "purification of type" (elimination of deformed individuals that departed too far from an immutable ideal morphology). But, for Darwinians, natural selection is not merely an executioner, but a creative craftsman as well. By accumulating favorable variation, bit by bit and generation by generation, natural selection superintends the process of organic change.[85]

The time scale involved is the feature distinguishing natural selection generally from Darwinian natural selection. As indicated, the crux of the problem of reconciliation is the differing rates of appearance. If physical evolution is gradual while the soul emerges fully formed, it can be difficult to relate one to the other. From among possible evolutionary scenarios, then, Darwinism should be most at odds with

---

[85]Ibid., 583.

soul models. But one can be a complete natural selection evolutionist who eschews gradualism. This alternative has evolutionary changes happening suddenly, rather than in gradual steps. The challenge of overlaying a self-preserving soul onto a biologically evolving form is easier if both units appear in whole chunks. The next section explains why this alluring alternative is not available to the soul-theorist.

2. Gradualism versus punctuated equilibria

Thomas Huxley, Darwin's primary advocate in the public furor that would later erupt, regretted that Darwin had so committed himself to the Linnaean maxim "Natura non facit saltum" ("Nature does not make leaps").[86] The rate of evolutionary change continues to be a topic of heated debate in scientific circles.

The nub of the controversy is the fossil record. It does not support Darwin's model of gradual, species-wide transformation, and he knew it. The paleontological record presented two problems: First, it failed to offer a single example of the kind of infinitely graded transformation he imagined. Steps were evidenced, but with no intermediary connecting forms. Second, species which endured millions of years in the fossil record demonstrated very little gradual change in all that time. In other words, stasis, and not gradual change, is the rule for established species. His response was to devote many pages of the *Origin* to explain the

---

[86]Gould, *The Panda's Thumb*, 179.

unavoidable incompleteness of the fossil record. A complete record, he believed, would in fact support his model.

In the 1970's a contrary position crystallized around a paper written by Niles Eldredge and Stephen Jay Gould.[87] They suggested that the fossil record be taken on its face, and that Darwinian-type species transformation was a rare event. More commonly, speciation events occurred rapidly, and in isolated populations. They termed their non-Darwinian model "Punctuated Equilibria" [P.E.]. "At its simplest, punctuated equilibria entails the recognition of stasis and the realization that patterns of change in the fossil record are best explained by allopatric speciation."[88]

> The theory of allopatric (or geographic) speciation suggests a different interpretation of paleontological data. If new species arise very rapidly in small, peripherally isolated local populations, then the great expectation of insensibly graded fossil sequences is a chimera. A new species does not evolve in the area of its ancestors; it does not arise from the slow transformation of all its forebears. Many breaks in the fossil record are real.[89]

---

[87]Niles Eldredge and Stephen Jay Gould, "Punctuated Equilibria: An Alternative to Phyletic Gradualism," in *Models in Paleobiology*, T.J.M Schopf (ed.), (San Francisco: Freeman, 1972), reprinted in Niles Eldredge, *Time Frames: The Rethinking of Darwinian Evolution and the Theory of Punctuated Equilibria* (New York: Simon and Schuster, 1985), 193-223.

[88]Eldredge, *Time Frames*, 120.

[89]Eldredge and Gould, "Punctuated Equilibria," 193.

Branching, and not transformation, was the central image in speciation. And speciation, when it occurred in isolated, peripheral populations, was rapid. In fact, whereas the traditional model had speciation as the *result* of evolution, an argument could be made that it was in fact the *precondition* of major evolutionary changes.

Contrary to the idea that variation arises in infinitesimal grades which natural selection then husbanded for the creation of new traits, variation was held to be prolific but masked until a new environmental context allowed it to be expressed. With consistent selection pressures in a novel environment, the new form would become quickly fixed, yielding a new species which may eventually spread beyond its original confines to compete with the parental species.

A key idea which enables the model of punctuated equilibria to work was not available to Darwin:[90] the genetic theory of inheritance. Significant change can be effected very quickly in either of two ways. First, genes can be either "structural," or determining the development of physical traits, or "regulatory," which determines the functioning of other genes. "The genetic command to grow at a given time, at a given rate,

---

[90]Although a case could be made that Darwin anticipated many of the primary points of punctuated equilibria, particularly the importance of geographic isolation for the creation of new species (cf. *The Origin of Species*, 94, 148, 150) and the origin of fossil species in locations other than that where their remains are first identified (302). Had he possessed a particulate theory of genetic inheritance, he might have found his way to this result.

and in a given way is easily stifled or modified, and the result may be dramatic."[91] The difference between the bear and the giant panda, for example, "could have been achieved by a very small number of regulatory genetic changes — perhaps just two or three."[92] Second, very minor changes can effectively isolate a population reproductively speaking. In species with restricted mating periods, for example, a slight shift in timing can cause two co-existing populations to become reproductively isolated.

This debate about evolutionary tempo is relevant for us in the following way. The soul-theorist could concede that the Darwinian gradualist evolutionary model does present intractable intellectual problems. This obstacle could possibly be overcome if the evolutionary appearance of the physical body were chunked in the way the soul is taken to be. If the anatomical emergence of *Homo* happened rapidly instead of a series of infinitely graded steps, it becomes reasonable to at least posit that the soul emerged similarly, thereby minimizing the conflict between the two theories. The point of conflict, in other words, is not with evolutionary theory, but with Darwinian gradualism.

Soul-theorists would be aided however only if the anatomical "jumps" were truly instantaneous, that is, occurred in one generation. At

---

[91]Stevens, *The New Evolutionary Timetable*, 130.

[92]Ibid., 129.

one time such changes were considered feasible and fell under the rubric of saltationism. Indeed, saltatory evolution has been proposed in the past precisely for the purpose "to defend their belief in the human soul."[93] One example of saltationism is Richard Goldschmidt's "hopeful monsters" who sprung fully formed from primitive parents: Rather than being bridged by developmental gradations, "the first bird hatched from a reptilian egg"[94] in one fell swoop.

The problems with this kind of theory are immediately obvious. Even if it happened, it would have to happen repeatedly, simultaneously, so as to provide reproductive mates for the new form: "With whom shall Athena born from Zeus's brow mate? All her relatives are members of another species."[95] The jump to a new species cannot therefore be made in a single generation, and this is the only kind of nongradual evolution which would aid soul-theorists.

P.E. uses much the same language as a more strictly saltationist position, and can therefore seem to offer succor to the soul-theorist, but in fact it is a much different theory.

> For some researchers, the picture of evolutionary change is one of long periods during which individual species remain virtually unchanged, punctuated by abrupt events of change when a descendent species arises. This model of

---

[93]Bowler, *Theories*, 49.

[94]Quoted by Hanegraaff, *The FACE,* 41.

[95]Gould, *Panda's Thumb*, 190-191.

> evolutionary change is called punctuated equilibria, which refers to the fact that in the evolutionary history of some groups long periods of stasis (relative stability) are broken or punctuated by short bursts of evolutionary change stimulated by speciation.[96]

"Short" is here relative to geologic time, and not a description of the real time, as required by the soul-theorist.

> By "rapid" we do not mean overnight or even in a decade but rapid relative to the average geological duration of a successful species. This duration is sometimes measured in hundreds of thousands of years, but usually in millions of years. In this context, a peripheral isolate that takes a few hundred years or even a thousand years to speciate has done so in a fraction of 1% of the total time it is likely to endure as a species before extinction.[97]

This process may be "rapid" by geologic standards, but it is still too slow to aid soul-theorists.

Even within P.E. theory anatomical precursors do arise, but in geographically isolated quick increments, not in the species-wide infinitesimal gradations supposed by the Modern Synthesis. So evolutionary change is still incremental, although in P.E. "anatomical change would be complete in ten or one hundred generations, depending on a particular species' generation time"[98] and not as slow and gradual

---

[96]Frank E. Poirier, *Understanding Human Evolution*, Third Edition (Englewood Cliffs, NJ: Prentice Hall, 1993), 27.

[97]Luria, Gould, and Singer, *A View of Life*, 624.

[98]Poirier, *Understanding Human Evolution*, 28. Significant changes can arise in even less time than Poirier suggests. A recent study examined a species of fruit fly (*Dropsophila subobscura*) introduced into

as Darwin had imagined. Importantly, P.E. does not posit any causal agents. Instead it only claims that the accepted components of the Synthesis are most operative in reproductively isolated populations, rather than species-wide as Darwin assumed.

> We [Gould and Eldredge] do not advance some special theory for long times and large transitions, fundamentally opposed to the processes of microevolution. Rather, we maintain that nature is organized hierarchically and that no smooth continuum leads across levels. We may attain a unified theory of process, but the processes work differently at different levels and we cannot extrapolate from one level to encompass all events at the next.[99]

This claim marks a significant reconceptualization of evolutionary theory, but not one which will help the soul-theorist.

---

the United states in 1978 into the Pacific Northwest of the United States. While a study just ten years ago "failed to find any size differences in the immigrant flies," today they have "evolved a pattern of size differences that mirrors the pattern in their European forebears" (that is, the new species in the north has longer wings (up to 4% among females) than their relatives to the south). These results show "that natural selection is not only at work but working at breathtaking speed." Tim Appenzeller, "Darwin, as Fit as Ever," *U.S. News & World Report,* January 24, 2000, 49; "Evolution May Not Be Slow or Random," *Science News*, January 29, 2000, 75.

Of similar importance is not the mere fact of this change, but that it is a change in conformance with Bergmann's rule, "A general pattern of geographic variation within a species leading to larger sizes in colder regions." Luria, Gould, and Singer, *A View of Life*, 761. This conformity demonstrates the predictive power of evolutionary theory.

[99]Stephen Jay Gould, "G.G. Simpson, Paleontology, and the Modern Synthesis," in *The Evolutionary Synthesis*, edited by Ernst Mayr and William B. Provine. Cambridge: Harvard University Press, 1980, 170.

While the debate between the two kinds of evolutionary pace continues, the outcome fails to inform this discussion because neither allows for the instantaneous appearance of new species, the only kind of categorical change that would comfortably comport with the appearance of a soul in toto. Whether conceived in terms of conventional Darwinism (slow and gradual) or P.E. ("jerky"), the proponent of the soul must still argue that the soul developed (regardless of the scale of time) rather than suddenly appeared.[100] So while important technical distinctions remain between the terms "natural selection" and "Darwinism," these

---

[100]Another alternative strategy for soul-theorists will not be discussed here, largely because it is similarly a dead-end. Perhaps, they might argue, the incrementally developed anatomical traits of humans allowed for the *emergence* "all at once" of the soul. This possibility is interesting only if "soul" is here taken to be synonymous with "mind," and indeed there are arguments which relate the mind/soul to the body in this way. Cf. Jaegwon Kim, *Philosophy of Mind* (Boulder, CO: Westview Press, 1998).

But as Kim points out, emergentism is a form of nonreductive physicalism (pp. 226-229). While soul-theorists would appreciate the nonreductive features of emergentism, meaning that mind/soul is not understandable in terms of its physical components, the theory still requires an ontological physicalism, meaning that it adheres to a principle that "All that exists in the space-time world are the basic particles recognized in physics and their aggregates" (p. 227). The soul-theorist is still left with the burden of identifying (with some measure of useful specificity) the physical foundation supporting this emergence. It is impossibly difficult to conceive how a mind/soul which preserves the self after death can adhere to such a physicalist premise.

Another alternative we should at least recognize is the argument that the soul preexisted, and that what "emerged" or "evolved" was the mechanism by which the soul bonds or interacts with body. It is unclear that this reformulation does anything more than alter the terms of the problem, leaving unaffected the problem itself.

distinctions are irrelevant for us and consequently the words will be used interchangeably.

E. The Self Preserved by the Soul

The position adopted in this thesis is that the only concept of the soul which is *interesting* is one which preserves the self. It is very possible that something other than this is true about the soul — e.g., that it does not exist at all, or that it preserves something other than the self. What is assumed is that if it were known that the soul were other than self-preserving, the soul would cease to be of first-order philosophical interest to the degree that it historically has been, although it could well be an important means to some other intellectual end.

What does it mean to speak of a "self," which task it is for the soul to preserve? Here appeal is made only to our pretheoretical intuitions. The self is that toward which we gesture with such terms as "I", "Me," "My" and "Mine." The self is what makes me *me* and not you, but which also makes the me of today the same as the me of twenty years ago. It is not important here to speculate as to how these intuitions arise, or of what they are constituted, or even if they are valid.[101] It is asserted as

---

[101]The neo-Freudian Henry Stack Sullivan, for example, denies the reality of personal individuality:

> One of the greatest difficulties encountered in bringing about favorable change is this almost inescapable illusion that there is a perduring, unique, simple existent self, called variously "me" or "I," and in some strange fashion, the

brute fact that the experience of a self is recognizable by all normal adults, despite these lingering problems.[102]

Galen Strawson, no friend to the concept of the self, does sketch more technically the presumed elements of the common intuitions about the self:

> I propose that the mental self is conceived or experienced as *(a)* a *thing*, in some reasonably robust sense. It is thought of as *(b) mental*, in some sense. It is thought of as *single*, in a way that requires further specification: both when considered *(c)* at a time and *(d)* through time. It is thought of as something that is *(e) ontologically distinct* from its thoughts, experiences, and so on, and indeed from all other things. It is thought of as something that is *(f)* a *subject of*

> patient's, or the subject person's, private property.... It makes no sense to think of ourselves as "individual," "separate," capable of anything like definitive description in isolation.

Sullivan, "The Illusion of Personal Individuality," 329. He here actually confuses two separate claims: It is not fatal to the idea of a "self" to concede that that self is not "capable of anything like definitive description in isolation." This latter claim is nothing but ordinary structuralism. The self may not be its own explanation (e.g., it has a social history, and thus one must look external to the self to explain much of its contents), but that does not impact its genuineness.

Simpson takes exactly the opposite view from Sullivan. For him, the hallmark of human evolution has been the (a) progress in awareness and reactibility to environmental stimuli, which (b) increases the range of reactions by each organism, resulting in (c) an accompanying progress in individualization. "Individualization is a prerequisite for the human type of socialization." *Meaning of Evolution*, 261.

What for Sullivan is at best a misdirecting epiphenomenon, if not outright illusory, is for Simpson the very real evolutionary achievement which sets humans apart from the rest of the animal kingdom.

[102]For a thorough consideration of the problem of the self, see Parfit, *Reasons and Persons*, Part Three.

> *experience*, a conscious feeler, thinker, chooser, decider, it is thought of as *(g)* an *agent*, and it is thought of as something that has a certain character or *(h) personality*.[103]

These criteria seem to me sensible.

We are very emotionally attached to our own selves, again independently of whether this attachment is rationally justified. The self functions as the prototype for all our other knowledge about the rest of the world. We understand other people, for example, only by analogy with our own psychological motivations. If it is difficult to conceive of the self without a world, it is impossible to imagine the world without the self. At the very least, the self is always the one looking at that world, even in the imagination. For each of us, our own self is the only necessarily real thing (or so claimed Descartes).

As such, we struggle hard to preserve it in this life. We would prefer not to believe that these efforts are ultimately futile.[104] Perhaps

---

[103]Galen Strawson, "The Sense of the Self," in *From Soul to Self*, edited by M. James C. Crabbe (London: Routledge, 1999), 132. Given his intention to refute the concept of the self, this characterization is admirably dispassionate.

[104]David Darling gets it completely backwards when he argues that the way to rob death of its sting is to surrender the idea of the "I". Without "the illusion of the self" "death is deprived of its victim, so that the basis for fear and sorrow of death is undermined." David Darling, *Soul Search: A Scientist Explores the Afterlife* (New York: Villard Books, 1995), 171. "Only when there is no self left is there no one who can die" (p. 177).

Strictly speaking, Darling is correct. But he has "solved" the problem of death in a bizarre way. He makes avoiding death-fear more important than preserving the self, when for most persons the priorities

due to the inertia of constantly willing the self to continue, or perhaps arising from the fact that when we imagine death, we imagine ourselves there, observing it, but for whatever cause we want (desire, desperately crave) that our selves survive beyond the death event.

What does it mean to "survive" death? John Perry neatly captures the important sense involved:

> Survival means that tomorrow, or sometime in the future, there will be someone who will experience, who will see and touch and smell — or at the very least, think and reason and remember. And this person will be *me*. This person will be related to me in such a way that it is correct for me to anticipate, to look forward to, those future experiences. And I am related to her in such a way that it will be right for her to remember what I have thought and done, to feel remorse for what I have done wrong, and pride in what I have done right. And the only relation that supports anticipation and memory in this way, is simply *identity*.[105]

---

are reversed: Our first goal is to preserve the self, and only secondarily to mitigate the death anxieties which arise in the wake of this quest of self-preservation.

More pointedly, scholars who have reached the conclusion that there is no "self," just as Darling desires, also deny that this would alleviate the problem about death, as Darling hopes. Having articulated his reasons why there is no self (no "Me*"), Galen Strawson surprises his readers:

> Now for an exception. You might expect me to say that when I think of my death at some unspecified future time, I think that it is not Me* who is going to die, or at least that I do not think that it is Me*. But I do think that it is Me* that is going to die, and I feel fear of death. It's only when I consider future events *in life* that I do not think it's Me*.

Strawson, "The Sense of the Self," 142.

[105]John Perry, *A Dialogue on Personal Identity and Immortality* (Indianapolis, IN: Hackett, 1978), 3-4.

This description applies to survival in the sense of any continuity of the self, even while living.[106] In the event of amnesia, for example, the body has not died but the self has been extinguished; in this instance there has been no survival even though there has been no loss of life.

Our requirement is that survival as here described pertains to death. This is the nonspecific sense in which we are requiring the soul to be self-preserving.

For each theory reviewed in Part II, the question is asked: For the kind of soul being defended, is it such that it could contain what we intuitively take to be our "self"? Note that we are not requiring that the soul-theory *actually* contain this concept. Most will not be structured in these explicit terms. More commonly, we will be looking for any premise

---

[106]Thomas Nagel denies that any characterization such as Perry's adequately captures "the conditions that must obtain if two experiential episodes separated in time are to belong to a single person," whether or not the episodes are separated by death. Thomas Nagel, *Mortal Questions* (Cambridge: Cambridge University Press, 1979), 199. For him, technical scrutiny of the idea causes the "self that appears to the subject to disappear" (p. 201). He may be right that the question of identity "cannot be exhausted by any account in terms of memory, similarity of character" (p. 200). For him these qualities are "never sufficient," and probably not even necessary. I would argue that the variables included by Perry may be found by formal analysis not to be sufficient, but they are surely necessary for any intuitively sensible conceptualization of what it means to have a self, and for that self to extend over time and beyond death. Perry's depiction, then, *contra* Nagel, can be read as what must be minimally true if the claim is to be valid.

of the theory which precludes this understanding of the self. What, in other words, is being argued to survive death?

F. The Criterion for Compatibility with Evolution

We have seen, then, that theorists who choose to defend the theory of a self-preserving soul on rational grounds cannot avoid conflict with evolutionary theory, and that it is the burden of the former to conform to the strictures of the latter. It remains to clarify more precisely the terms of that conformity. What must a soul-theory contain to be deemed "compatible with evolution"?

Many theories are to be evaluated by this criterion. It must therefore be specific enough to answer our philosophical question, yet general enough so that soul-theories formulated prior to the *Origin of Species* are likely to address the relevant issues. For this reason we cannot expect to employ so restricted a criterion as whether or not the soul-theory resolves the sorites paradox of soul injection. A weaker standard is required, but one which can be widely applied. Consequently the criterion for evolutionary compatibility shall be whether or not the soul-theory allows for the principle of Psychological Continuity.

Darwin argued that just as anatomy supports the thesis of a common descent for animals, so to does an examination of the respective mental lives.

> It has, I think, now been shewn that man and the higher animals, especially the Primates, have some few instincts in

> common. All have the same senses, intuitions and sensations — similar passions, affections, and emotions, even the more complex ones; they feel wonder and curiosity; they possess the same faculties of imitation, attention, memory, imagination, and reason, though in very different degrees.[107]

The most significant difference is acknowledged to be that of the moral sense.[108] But careful ethological study shows that animals are not utterly devoid of even this higher capacity:

> Besides love and sympathy, animals exhibit other qualities which in us would be called moral; and I agree with Agassiz that dogs possess something very like a conscience. They certainly possess some power of self-command, and this does not appear to be wholly the result of fear.[109]

With no exception, then, "the difference in mind between and the higher animals, great as it is, is certainly one of degree and not of kind."[110]

---

[107]Darwin, *The Descent of Man,* Vol. I, 48-49. Although he did not discuss the topic in any detail, Darwin realized even in his earlier *Origin* that his claims about organic structural evolution probably applied to mental traits as well: "the canon in natural history, of 'natura non facit saltum' is applicable to instincts as well as to corporeal structure." Darwin, *Origin of Species*, 263.

Contemporary research has shown that in the details Darwin's portrait of animal-human similarities was optimistic. For example, "Despite earlier claims, there is now considerable doubt about the extent to which any monkeys or apes have been shown to *imitate* behaviour outside a context of social interaction with humans." Noble and Davidson, *Human Evolution, Language and Mind*, 42.

[108]Darwin, *The Descent of Man,* Vol. I, 70.

[109]Ibid., 78.

[110]Ibid., 105. Scattered through this text are examples of Darwin's anthropomorphic characterization of other species. In Volume II, for example, birds "unconsciously prefer" (p. 50), "apparently enjoy" (p. 51),

Only if this claim is true can we hope to use imaginative empathy to understand the reasons and motivations behind observed animal behaviors.[111] Darwin presented his argument for this hypothesis in his *The Expression of the Emotions in Man and Animals*, wherein he argued

> That our expressions of emotion are universal (that is, innate not learned) and the product of our evolution. Neither our expressions nor our emotions are unique to human beings; other animals have some of the same emotions, and some of the same expressions shown by animals resemble our own.[112]

According to primatologist Suzanne Chevalier-Skolnikoff, "Darwin's central hypothesis — that the facial expressions of non-human primates and man are similar — has been strikingly confirmed, strongly

---

"take pleasure in practising whatever instinct they follow at other time for some real good" (p. 54), display behaviors just short of "conscious vanity" (p. 96), can regard affairs "as of the highest importance" (p. 102), "exhibit benevolent feelings" (p. 109), "evinced unbounded curiosity, and clearly had 'the idea of property and possession'" (p. 109), and can feel jealousy (p. 110).

In *Expression* Darwin claims that "Even insects express anger, terror, jealousy, and love by their stridulation" (p. 347).

[111]Ibid., 122.

[112]Charles Darwin, *The Expression of the Emotions in Man and Animals*, Paul Ekman (ed.), (New York: Oxford University Press, 1998), xxii.

This volume had also another goal besides that of demonstrating the continuity between man and animals. It also endeavored to argue that since all humans express the same emotions, the races diverged after full human status had been achieved and not, as some supposed, earlier.

supporting his theory that human facial expressions have evolved from those of man's non human primate ancestors."[113]

This conclusion, that the gap between human and animal mind is one of degree and not of kind, is the core of the thesis of psychological continuity.[114] This principle holds that "psychological acts, states, and functions in lower animals *model* those in higher animals" in that the former is "the very same kind of things as the latter, a similar kind of thing, or at least something analogous to the latter."[115]

Psychological continuity can be contrasted with a belief that humans and animals are different in kind, not degree, that humans possess something qualitatively distinct from anything any animal possesses in any way. This position is neither irrational nor indefensible, but it is difficult if not impossible to reconcile with a prior understanding of *Homo sapiens* as evolutionarily connected with other animal species.[116] These historical developmental linkages strongly imply that

---

[113]Ibid., 144.

[114]Parfit (*Reasons and Persons*, 207) means something completely different when he uses this term.

[115]Gary Matthews, "Augustine and Descartes on the Souls of Animals," in *From Soul to Self*, edited by M. James C. Crabbe (London: Routledge, 1999), 95.

[116]This is not to say that there is *nothing* distinctive about human beings, and many writers work diligently to find just such a marker. Darwin's own conclusion was that only humans blush because this reaction requires "a sensitive regard for the opinion, more particularly for

different species will bear the marks of this relatedness. We expect, in other words, that even distantly related animals will be at least analogous on the most fundamental psychological properties. In some cases this expectation may be frustrated if the species are too distantly related or if the analyzed properties are less fundamental and more specialized, but nevertheless it should be our default presumption.

A soul-theory which denies the principle of psychological continuity is one which is therefore immediately incompatible with evolutionary theory. For our purposes, therefore, a soul-theory shall be deemed presumptively "compatible with evolution" if it allows for the principle of psychological continuity. We can adopt this position even realizing that a soul-theory which allows for psychological continuity can still be incompatible with evolution, but at deeper levels than those examined here.

G. Project Outline

Part II of this thesis reviews several attempts to defend the existence of the soul, and judges them according to three criteria:

First, is the defense internally consistent; that is, is it a good defense when judged by its own standards?

Second, is the kind of soul defended by this argument one which adequately contains that which we designate the "self"? And

---

the depreciation of others, primarily in relation to our personal

Three, does the defense allow, either explicitly or by implication, for the principle of psychological continuity?

---

appearance, especially of our faces." Darwin, *Expression*, 334.

PART II

PHILOSOPHICAL TEXTS

CHAPTER II

PLATO'S
PHAEDO (c. 387-380 B.C.E.)

The soul is a pervasive concern in the dialogues of Plato. *That* it is he always takes for granted, never feeling it necessary to argue for the soul's existence. Not only is Plato not a materialist, he seems not to have regarded the materialist position worthy of extensive rebuttal.[1] His concerns instead are typically to ponder what the soul is like, and how it is best nurtured.

That the soul possesses the virtue of being immortal, however, is a claim he defends at least twice.[2] Briefly and almost incidentally in Book

---

[1]Note the refusal to allow for a purely materialistic understanding of death in the following claim by Socrates:

> And by death do we not mean simply the departure of soul from body? Being dead consists, does it not, in the body having been parted from the soul and come to be by itself, and in the soul having been parted from the body, and being by itself. Can death possibly be anything other than that? [To which Simmias replies:] No, it can only be that. (64c)

R. Hackforth, *Plato's Phaedo* (Cambridge: Cambridge University Press, 1955), 44. All quoted passages from this dialogue come from this translation.

Plato does later address a particular kind of materialism when he considers the possibility that the soul is an "attunement of the bodily constituents" (85b-95a). It is unclear, however, that this argument precludes other forms of materialism.

[2]Hackforth (*Plato's Phaedo*, 19-24) offers the following list of dialogues in addition to the *Phaedo* which include arguments for immortality: *Meno, Symposium, Republic, Phaedrus,* and the *Laws, Book*

X of the *Republic*, Plato argues that "our soul is immortal and never destroyed."[3] He reasoned that everything has a "good and a bad," and that a thing can only be destroyed by the evil peculiar to it. Examples include blight for grain, rot for wood, and rust for iron or bronze. The evils of the soul — "Injustice, licentiousness, cowardice, and lack of learning"[4] — warp the soul, but do not "make it waste away until, having brought it to the point of death, they separate it from the body."[5]

Ten. The argument presented above addresses the *Republic*'s arguments about destruction by the proper evils, and also shows that the *Meno* is incorrectly included by Hackforth: that dialogue assumes the soul's immortality; it does not defend this assumption.

The remaining dialogues also fall outside the scope of this thesis. The *Symposium*, according to Hackforth, offers only "a vicarious survival, not an immortality of the personal self, the individual soul, that is proclaimed" (p. 20). In a similar vein, T.M. Robinson concludes that "One of the most condensed and abstruse arguments for the soul's immortality is to be found at *Phaedrus* 245c-246a. It is not even clear whether 'soul' here is meant to refer to soul collectively [i.e., it could be a defense of the 'World Soul' described by the *Timaeus*] or to individual souls." T.M. Robinson, *Plato's Psychology* (Toronto: University of Toronto Press, 1970), 111.

Finally, "the soul of *Laws* X seems to be a cosmic version of the individual soul seen as a *life*-soul, just as the individual soul seen in its *noetic* or cognitive aspect had its cosmic equivalent in the rational World Soul of the *Timaeus* and *Politicus*." Robinson, *Plato's Psychology*, 153.

All of these other defenses of the soul, therefore, fall short of the present concern to examine Plato's defense of the personal, self-preserving soul of the individual.

[3]Plato, *Republic*, translated by G.M.A. Grube (Indianapolis: Hackett Publishing, 1992), 608d. All quoted passages from this dialogue come from this edition.

[4]Ibid., 609b.

[5]Ibid., 609d.

> If the soul's own evil and badness isn't enough to kill and destroy it, an evil appointed for the destruction of something else will hardly kill it.... If the soul isn't destroyed by a single evil, whether its own or something else's, then clearly it must always be. And if it always is, it is immortal."[6]

The immortality of the soul is far from being Plato's primary concern in the *Republic*, which is instead to argue that being virtuous is best for its own sake and not for any rewards it might bring. That includes postmortem rewards, so it not entirely clear that arguing for the soul's immortality doesn't subtly undermine his larger purpose in that dialogue. If such rewards exist, it is hard to imagine that they are not sufficient reason for being virtuous. The argument that we should practice virtue for its own rewards is intellectually interesting, but adds nothing more to our motivation to behavior virtuously.

In any event, this is a particularly unconvincing defense of the immortality of the soul because of its structure. Plato essentially argues that X is hurt only by those things that hurt X; he offers no *a priori* reason to think X might be hurt by something. The examples of the *Republic* gives us specific harms, but no discussion of how each item on the list effects those harms. We are given, in other words, a list of particulars when what we need is an identification of principles.

Blight for grain and rot for wood, for example, are both examples of fungal destruction. Rather than listing the particulars, we could extract

---

[6]Ibid., 610e.

a general destructive principle that anything which consumes the matter destroys it. From this we could learn whether something can destroy X not by referencing a list of particulars but by studying the agent to see if it possesses the requisite property. This step, however, Plato does not take. However, a list of particulars cannot make the logical leap from the identification of the things *known* to hurt X to the conclusion that if none of these things can destroy it, nothing can.

The serious work of presenting a reasoned defense of the immortality of the soul is done in another dialogue, the *Phaedo*.[7]

---

[7]We can take for granted how influential this argument will prove in Western history. David Gallop suggests that it "has become almost a charter document for a certain world view, which may conveniently be dubbed 'dualistic'." David Gallop, "Introduction," *Phaedo* (Oxford: Oxford University Press, 1993), ix. "Its teaching largely anticipates the Christian conception of the soul as a spiritual substance, conjoined with the body during life, separated from it at death, and capable of existing thereafter in a disembodied state for all eternity" (p. xv).

The influence of the Platonic conception of the soul on Christian theology has in fact recently been the subject of discussion on Hell. The issue seems to be whether those suffering in Hell do so for all eternity, or are "annihilated."

> Proponents of this theory, called "annihilationism," argue that the traditional belief in unending torment is based more on pagan philosophy than on a correct understanding of the Scripture.... 'How can Christians possibly project a deity of such cruelty and vindictiveness' as to inflict 'ever-lasting torture upon his creatures, however sinful they may have been? asks [Clark H.] Pinnock.... [John R.W.] Stott observes that in biblical imagery, the fire's main function is to destroy and that while the fire of hell may be eternal and unquenchable, "it would be very odd if what is thrown into it proves indestructible." And [Philip E.] Hughes argues that the traditional belief in unending punishment is linked to the

A. Criterion One: Theory Consistency

The *Phaedo* is the dialogue set immediately preceding Socrates' death. Its dramatic purpose is to allay the fears of his friends concerning his fate: "I am the less disposed to complain, but am of good hope that there is a future for those who have died."[8] Indeed, he believes that "all those who betake themselves to philosophy in the right way [namely, himself] are engaged in one thing only, namely training themselves for dying and being dead."[9] A true philosopher is less attached to his body, less prone to its tyrannical, distorting desires. He lives for the improvement of his soul, and therefore "One who pays no regard to the pleasures which come by way of the body has, they would say, one foot in the grave."[10]

This conclusion in the mouths of the *hoi polloi* could be an insult: Anyone who does not indulge in the appetites of the body is a dull,

---

> Greek notion of the innate immortality of the soul — a belief he says is based more on Plato than on the Bible. "The immortality of which the Christian is assured is not inherent in himself or in his soul but is bestowed by God," says Hughes.

Jeffrey L. Sheler, "Hell Hath No Fury," *U.S. News & World Report*, January 31, 2000, 50.

[8]Plato, *Phaedo*, 63c.

[9]Ibid., 64a.

[10]Ibid., 65a.

affectless person little better than a corpse, hardly someone to be emulated. But the philosopher knows the directive as a literal truth: By ignoring his body and exercising his soul, he has already gone as far in this life as one can toward effectuating the liberation of the soul at death, and any movement in that direction a wondrous achievement.

All good and well, but this depiction does not qualify as a proof that it is true. Although the existence of the soul is never in doubt, there is nothing in that concept which necessitates its immortality, or, if being immortal, that that immortal soul is "me." That task, then, is to provide a rational foundation for this cheery picture of a death-surviving soul, with special postmortem rewards for philosophers.[11]

---

[11]The belief in afterlife rewards was uncommon among Athenians: "the expectation of rewards for virtue in the afterlife was not only very uncertain, but also very uncommon." Jon D. Mikalson, *Athenian Popular Religion* (Chapel Hill, NC: University of North Carolina, 1983), 80. Michael L. Morgan labels this the "Delphic theology," referring to the "attitude of separation between the divine and the human, of discontinuity, of human limits and hence of the temptation to illicit self-esteem and pride (*hubris*)." Michael L. Morgan, "Plato and Greek Religion," in *The Cambridge Companion to Plato*, edited by Richard Kraut (Cambridge: Cambridge University Press, 1992), 231.

The contrast is with the "Orphic-Bacchic-Pythagorean-Eleusinian world"

> which proclaimed a theory, unfamiliar to native Greek mythology and religion, that the soul of man is divine and of divine origin; that the body is its impure prison-house where it is in danger of contracting stain; that by elaborate abstinences the soul might retain its purity, and by sacramental and magic methods the pure soul might enjoy in this life and in the next full communion with God."

Toward this end Socrates offers four "proofs" for the immortality of the soul: (1) the Opposition argument, (2) the Recollection argument, and the arguments from (3) Forms and (4) Entailment.

1. The opposition argument (69e-72d)

According to Socrates, of necessity any one of a pair of opposed qualities emerges from the other in a cycle of transformation. "Alive" and "dead" are just such a pair of opposites, and as we already know that those things which are dead come into being from those which are alive, so too do those things which are alive come into being from those which are dead. This at least proves the pre-existence of the soul before its current incorporation (*metensomatosis*):

> 'Then what is it that comes to be from that which is alive?' 'What is dead.'
> 'And from what is dead?' 'I am bound to admit that the answer must be "What is alive".'
> 'Then, Cebes, living things and living people come into being from dead things and dead people.' 'Evidently.'
> 'Hence our souls do exist in Hades?' 'Apparently.'[12]

L.R. Farnell, *Outline History of Greek Religion* (Chicago, IL: Ares Publishers, 1974), 82-83. What Plato brings that is new to this Orphism is the replacement of "the emotional character of the ritual process with cognitive content." Morgan, "Plato and Greek Religion," 232.

In this vein it can be helpful to remember what is original with Plato and what he receives but perfects. For example, in addition to Plato redirecting the Orphic emphasis from the ecstatic to the cognitive, he also provides in the *Phaedo* "the first extensive discussion of the soul's immateriality" which was only implicit in Homer. Richard Sorabji, "Soul and Self in Ancient Philosophy," in *From Soul to Self*, edited by M. James C. Crabbe (London: Routledge, 1999), 10.

[12]Plato, *Phaedo*, 71d-e.

The persuasive potential of this argument depends upon the degree to which you would accept his premise that qualities emerge from their opposites. His own examples are not compelling: the transition of the bigger into the smaller, the stronger into the weaker, the slower to the faster, and the worse to the better, the more just from the more unjust. Two weaknesses are easily identified.

First, Socrates' illustrative instances are all comparative and quantitative (the less X to the more X) even though he intends to use the argument to support a qualitative opposition (the X to the not-X). Even if the argument as literally framed were acceptable, it is unclear how it extends to his intended application of the principle.

He pins his argument on the linguistic equation of dead/alive and small/big as both being oppositions of the same sort, when in fact they are not. All of his given examples are *contraries*, that is, they are opposites (like hot and cold) which admit innumerable gradations between them. But death and life have no intermediaries of this kind; they are contradictories. One is either dead, or one is not; one cannot be a little bit dead. Even when we speak of someone as being "barely alive," we mean that he is alive, but has approached the threshold to death; until that line is crossed, he is completely alive.

Obviously two different kinds of oppositions are involved here, contraries and contradictories.[13] The argument that one member of a pair necessarily "emerges from" the other is justified solely in terms of the former, but then applied to the latter. This application is unsupported. Because the change from smaller to bigger is thus of a completely different nature than the change from death to life, us must assign the argument from Opposition no persuasive power at all in this context.

The preceding paragraph assumed that the argument was valid at least for the examples he gives, questioning only his final application of the principle. In fact, it is doubtful that the broader Platonic corpus would support even that assumption.[14]

In the *Phaedo* Socrates requires that the change from smaller to bigger be an actual change, paralleling that from being dead to being alive. In the later dialogue the *Theaetetus*, Socrates questions this characterization: "Is it possible, Theaetetus, to become bigger or more in

---

[13]Small/big and dead/alive are different in another way as well. Whereas small/big are comparative relationships (an object is not inherently small or big, but rather it is these things relative to something else), being dead or alive is a property possessed of the object, regardless of any comparisons. All told, drawing conclusions about one drawn from properties of the other is a awkward project.

[14]"Never consider any one expression of Socrates' views in isolation from other expressions of Socrates' views." Terry Penner, "Socrates and the Early Dialogues," in *The Cambridge Companion to Plato*, edited by Richard Kraut (Cambridge: Cambridge University Press, 1992), 122.

number in any other way than by being increased?"[15] In that passage Socrates has already shown one way in which the answer to this question must be, "Yes":

> Here are six dice. Put four beside them, and they are more, we say than the four, that is, half as many again; but put twelve beside them, and we say they are less, that is, half the number.[16]

Examples such as this cause three statements "to fight one another in our souls": (a) "nothing can possibly become either greater or less, in bulk or in number, so long as it is equal to itself"; (b) "a thing to which nothing is added and from which nothing is taken away neither increases nor diminishes but remains equal"; and (c) "it is impossible that a thing should ever be what it was not before without having become and without any process of becoming."[17]

We should note that the dice example is only a problem because Socrates seems to treat "big" like "white." "White" is a quality (color) that the object possesses, at least by one interpretation. For any given object you can make property claims (X is white) or comparative claims (X is whiter than Y, that is, possesses the quality white in a purer form or

---

[15]Myles Burnyeat, *The Theaetetus of Plato*, translation by M.J. Levett (Indianapolis: Hackett, 1990), 154c. All quoted passages from this dialogue come from this edition.

[16]Ibid., 154c.

[17]Ibid., 155a-b.

greater quantity than does Y). "Big," unlike "white," can be used only in claims of evaluation (X is bigger than Y). Even when we seem to be using "big" to make property claims (X is big) we do so only because it is understood with the implication that 'X is big [relative to the normal value for that quality].' "Big" is not a property an object can possess in isolation.

At least in these dialogues[18] Plato does not mark this distinction between properties inherent to the object and those which are assigned

---

[18]By contrast, Plato successfully diffuses a related problem in the *Sophist*. There the difficulty is that

> For a statement, in order to be a statement at all, has to manage to say something, that is, there has to be something that gets said by it. But both in ordinary Greek and in the language of Greek philosophers a false statement is one that says what is not (or: what is not being). Yet what is not being does not seem to be something that is there to get said. Hence it would seem that there is nothing that gets said by a false statement. But in this case it fails to be a statement. So it seems that there can be no false statements.

Michael Frede, "Plato's *Sophist* on False Statements," in *The Cambridge Companion to Plato*, edited by Richard Kraut (Cambridge: Cambridge University Press, 1992), 397.

> Plato tries to explain where people got confused when they came to think that there is no such thing as what is not. The claim is that they came to think "what is not" or "not being" must refer to the contrary of what is. What is, is something of which it is true that it is *in some respect*. Instead of realizing that what is not, by contrast, is something of which it is not true that it is *in this respect*, they assumed that it is something of which it is not true that it is *in any respect*.

to it solely through contrast with other objects. He consequently treats "small" and "big" as actual properties, creating the tension he describes, especially from statement (c). There can be no "process of becoming" from "smaller" to "bigger" because the object (here, the six dice) are both simultaneously, relative to the appropriate comparative contexts. This is only a conceptual paradox if you treat "small" and "big" as properties, which would require you to claim that a single object simultaneously instantiates both members of an opposition.

Given Plato's understandings, the possible resolution examined in the *Theaetetus* conjoins Protagorean relativity with Heraclitean instability, a philosophical perspective which dissolves into a spectacular mess by the conclusion of the dialogue's second attempt to define "knowledge."[19] While the considered solution to the problem collapses, the question still remains. By his own example the correct response to

---

Ibid., 405 (emphases added). In other words, to say that "X is not Y" does not mean that X is the negation of Y, so that X is the not-beingness of Y. Rather, it means only that X is not *like* Y relative to some comparative criteria, and that both X and Y are there to be talked about.

We thus see Plato dissolving the paradox of the false statement by pointing out that many of our statements are evaluative and comparative, rather than substantive. This is the key insight which he lacks in the *Phaedo*, but one which if he had controlled it would have prevented him from asserting the argument from Opposition, at least as it is here cast.

[19]Plato, *Theaetetus*, 151d-186e. By the end of this examination, the combination of relativity and fluidity destroys the possibility of language. Protagorean epistemology leads to a loss of reference (171d); Heraclitean ontology leads to the loss of stable referents (183b).

the question will distinguish comparative changes from substantive transitions. However, in the *Phaedo* he treats comparative changes as qualitative changes, invalidating the point he needs to press.

The argument from Opposition, then, is invalid on two grounds. First, the examples he presents do not support the principle he wishes to establish. The change from smaller to bigger is a comparative, evaluative change which need not reflect any actual change in the object itself. Because there is no actual change, it makes no sense to claim that the smaller "must pass to being now smaller from being then bigger."[20] More importantly, even if the examples did establish the principle that everything comes into being from its opposite, he would have established that this is the case only for qualities ranging over a continuum (i.e., contraries), and not for contradictories of a present/absent nature.

---

[20]Plato, *Phaedo*, 71a. Plato suggests at 71b that without the principle of opposition, the world would experience a kind of terminal entropy:

> If this circular process of one opposite coming into being to balance the other were not always going on, if instead of that there were only a one-way process in a straight line, with no bending back, no turning in the other direction, you will realise that ultimately all things would arrive at the same state, would undergo the same experience, and the coming into being of things would be at an end.

There is at least one problem with this. It ignores the possibility that, although X does not come from not-X in the sense of the negation of X, X can come from not-X in the sense of not being X with respect to that quality, that is, from Y. Cf. Note 18 above.

Given these two weaknesses, we must conclude that the argument from Opposition fails.

2. The recollection argument (72e-77a)

In the *Meno* Socrates depicts the acquisition of knowledge as essentially a process of recollection (*anamnesis*). The context is to offer a resolution to "Meno's Paradox": "A man cannot search either for what he knows or for what he does not know. He cannot search for what he knows — since he knows it, there is no need to search — nor for what he does not know, for he does not know what to look for."[21]

Responding to this dilemma, Socrates begins by invokes the theological doctrine of the immortality of the soul relying solely on the authority of "wise men and women." He offers no independent support for the doctrine. From this belief he generates the following conclusion:

> *As the soul is immortal*, has been born often and has seen all things here and in the underworld, there is nothing which it has not learned; *so it is* in no way surprising that it can recollect the things it knew before, both about virtue and other things. As the whole of nature is akin, and the soul has learned everything, nothing prevents a man, after recalling one thing only — a process men call learning — discovering everything else for himself, if he is brave and does not tire of the search, for searching and learning are, as a whole, recollection.[22]

---

[21]Plato, *Meno*, translated by G.M.A. Grube (Indianapolis, IN: Hackett, 1976), 80e. All quoted passages from this dialogue come from this edition.

[22]Ibid., 81c-d (emphasis added). There seems to be a difficulty with this theory of Recollection which is not much discussed. That is, it accounts for our *current* condition: we have "been born often and [have]

As stated, the paradox would be valid, and the acquisition of knowledge impossible. It assumes, however, that the knowledge sought is something *new*. Socrates dissolves the paradox by claiming that each person already possesses all knowledge by virtue of his immortal soul. The challenge for the person is not to gain knowledge, but to recall such knowledge as he already has, that is, knowledge is internal and pre-existing (i.e., old). This task is difficult enough, but it does not present the inconsistencies thought to arise when the knowledge is assumed to be external and novel.

We should first note that immortality is not the attribute Plato needs attributed to the soul. The theory of recollection requires the soul's preexistence, which is not entailed by immortality. Ignoring that mix-up (Plato perhaps confuses immortal with eternal) we need to discern the role in the argument that the invocation of the soul's immortality is intended to provide.

Our question is, is this a *necessary* premise to ground the theory? To the extent that immortality provides the experiential background needed for recollection, yes, it is necessary. Without the mechanism of

---

seen all things here and in the underworld, there is nothing which [our souls have] not learned." It does not, however, tell us how this process might have been initiated. If all learning is a process of recollection, and if the information we are to recollect is not bestowed at one fell swoop, then how did we learn the new information we would later recollect?

immortality (or preexistence) the theory of recollection is only a metaphor. That Socrates proceeds to demonstrate recollection "on the hoof" by quizzing the slave boy clearly shows that Plato did not regard this theory as mere metaphor. Immortality, then, is the undefended theoretical ground for recollection.

As we have seen, in the *Meno* Socrates assumes an immortality of the soul to justify the theory of recollection. In the *Phaedo* he proceeds in the opposite direction, asserting the theory of recollection without justification to prove the immortality of the soul:

> If those objects exist which are always on our lips, a beautiful and a good and all reality of that sort, and if it is to that that we refer the content of our sense-perceptions, thereby recovering what was ours aforetime, and compare our percepts thereto, it must follow that as surely as those objects exist so surely do our souls exist before we were born; but if the former do not exist, all our argument will have gone for nothing.[23]

We are clearly meant to be already familiar with recollection. Cebes reminds Socrates that

> There is also another theory which, if true, points the same way, Socrates: the one that you are constantly asserting, namely that learning is really just recollection.... Hence we seem to have another indication that the soul is something immortal.[24]

---

[23]Plato, *Phaedo*, 76d-e.

[24]Ibid., 72e.

This prior familiarity — from the *Meno*, we can only assume — relieves Socrates of the responsibility of explicitly defending recollection in the *Phaedo*.

Were one reading only the *Phaedo*, one might conclude that the weakness of this argument from recollection is its admitted dependence upon the actuality of the realm of the Forms. The proof fails, we might respond, because no such realm exists.[25] But this would only make the argument erroneous, not inconsistent in the sense that it is self-contradictory.

But the really serious problem for Socrates comes in the circularity of his proof. The asserted immortality of the soul justifies the defended theory of recollection in the *Meno*, while the merely invoked theory of recollection justifies the defense of the immortality of the soul in the *Phaedo*. Within the Platonic corpus each claim is simply a restatement of the other.[26] That relationship may well be valid, but it is circular. What we require is an defense of the immortality of the soul in terms other than those whose own validity presumes the very fact we are trying to establish.

---

[25]Although Popper might disagree. The realm of the Forms bears resemblances to his Third World. Cf. Karl R. Popper, "On the Theory of the Objective Mind, In *Objective Knowledge* (Oxford: Oxford at the Clarendon Press, 1972), 153-190.

[26]The contentiousness of this argument is conceded. Although the textual support is to me obvious, not everyone will be convinced.

Given this circularity we must conclude that the argument from Recollection fails.

3. The formal argument (78b-80c)

The objection is made that even if the soul survives death, it may not do so for very long: "immediately upon [the soul's] departure, its exit, it is dispersed like breath or smoke, vanishing into thin air, and thereafter not existing anywhere at all."[27] Socrates' reply is that "compounded" or "composite" things are most likely to suffer the fate of dispersion into its parts; incomposite things having no parts cannot be similarly broken apart, and are thus immortal.

The body, he argues, is like the visible things, while the soul is like the invisible "unchanging objects [that] cannot be apprehended save by the mind's reasoning."[28]

> But when it investigates by itself alone, it passes to that other world of pure, everlasting, immortal, constant being, and by reason of its kinship thereto abides ever therewith, whensoever it has come to be by itself and is suffered to do so.[29]

---

[27]Plato, *Phaedo*, 70a.

[28]Ibid., 79a.

[29]Ibid., 79d.

Without invoking the doctrine explicitly, the gist of this defense is the Empedoclean principle that "like knows like";[30] that the soul is immortal is evidenced by the fact that it is able to know the Forms. The soul "has a far and away greater resemblance to everlasting, unchanging being than to its opposite."[31]

There are three problems with this argument. The first two involve the premises of the argument, that only incomposite things are immortal, and that the soul is incomposite. Other Platonic works from this period, especially the *Republic,* deny both of those claims.[32]

First, Plato backs off of this claim that only the incomposite can be immortal. In the *Republic* "Plato, having reconsidered the problem, has concluded (if only by implication) that a composite *can* very well be immortal, provided that its composition is of the 'best' sort."[33]
Second, the proof requires that the soul be "incomposite," and this is indeed the characterization presented in the *Phaedo*. But in the *Republic* the soul is tripartite, that is, it has three parts.

---

[30]Aristotle, "De Anima," translated by J.A. Smith, in *The Basic Works of Aristotle*, edited by Richard McKeon (New York: Random House, 1941), 410a27-29.

[31]Plato, *Phaedo*, 79e. This same reasoning is applied in the *Phaedrus* when it is argued that "it is proximity to the Ideas that makes the gods to be divine, that is, eternal." Robinson, *Plato's Psychology*, 101.

[32]See discussion of the *Republic* in this section's introduction.

[33]Robinson, *Plato's Psychology,* 53, referring to *Republic*, 611b5-7.

How to interpret this tripartiteness of the soul is contentious. The reading adopted here is that "parts" is here meant in the same way that the body has parts (organs). Other interpreters might regard the soul as homogenous, with different aspect or powers. The latter may be unpersuaded by the present argument, but there are at least three reasons why their position may need revision. First, the difficulties the Catholic Church has experienced in trying to make sense of the doctrine of the Trinity (three persons in one God) counsels caution in any attempt to defend distinct parts in a unitary entity.

Second, if the soul is homogenous but with different aspects or powers, the problem is raised about how they are actuated. Why *now* is the soul reasonable when *before* it was sensuous? A metapower is now required to differentially active the sundry parts of the unitary soul. Such solutions are possible,[34] but I find no indication that Plato was moving in this direction. The parts of the soul in the *Republic* are sovereign powers, not regulated by some unifying principle. This fact weakens the alternative interpretation of the tripartite soul as not consisting of parts in the ordinary sense of the word.

Finally, Socrates himself analogizes the parts of the soul with parts as ordinarily understood: "It isn't easy," he says, "for anything composed

---

[34]Aristotle's "active intellect" could be one attempt at such a metapower. Cf. Aristotle, "De Anima," Section 3.5.

of many parts to be immortal if it isn't put together in the finest way, yet this is how the soul now appeared to us."[35] The point of this passage seems not to be put a new spin on the word "parts," but to show that (contrary to the premises of the Formal argument) a composite thing as ordinarily understood can be immortal if its parts are "put together in the finest way."

Julia Annas agrees that "Plato never says that [the parts of the soul] are spatial or temporal parts."[36] In fact, she finds him "deliberately vague" on this point: "he is insisting that there is complexity in a single person without saying too much about how that complexity might be realized."[37] She does recognize, however, that in the *Timaeus* "he does talk as though the soul's parts are located in parts of the bodies, but there is no hint of that in the *Republic*."[38] So textually the best conclusion is that Plato does not *deny* that the parts of the soul are akin to parts of the body, but he does intimate that this is the case, in in other dialogues he is more forthright that this is the case.

---

[35]Plato, *Republic*, 611b.

[36]Julia Annas, *An Introduction to Plato's Republic* (Oxford: Clarendon Press, 1981), 124.

[37]Ibid.

[38]Ibid.

If true, that is, if the tripartite soul is reasonably understood as consisting of actual parts, then Plato has contradicted the initial premises of the Formal argument. Important premises of the Formal argument would therefore be inconsistent with the whole of the Platonic corpus and hence lacks suasive potency.

Finally, the argument accrues demonstrative power to the extent that the soul and the Forms are identical. Such identity cannot be the case, otherwise the soul would itself be a Form rather than a knower of the Forms. The dialogue concedes that the relationship is less than an equality:

> Would you say then, Cebes, that the result of our whole discussion amounts to this: on the one hand we have that which is divine, immortal, indestructible, of a single form, accessible to thought, ever constant and abiding true to itself; and *the soul is very like it*. (80b, emphasis added)

This concession is fatal to the argument. T.M. Robinson summarizes the dilemma thus:

> [The soul] cannot be set among the objects of the world of motion, for it is precisely these which, like body, are subject to change and decomposition. Yet it is apparently not an Idea; it is only "very like" the Ideas.[39] But if the argument is to succeed it must *be* an Idea. For only the reality of Ideas is

---

[39]This characterization inspires the question, If the soul belongs neither to the world of motion nor to the world of Forms, what world does it inhabit? I suspect Robinson has overstated the case here. The Platonic argument, as far as I can see, is not that the soul is impervious to *all* motion (i.e., change). It is altered, for instance, by experiences such as philosophic disciplines. The only claim is that one particular kind of motion (dissolution) is inapplicable to the soul.

> absolute, and any declination from this will involve a proportionate lessening in force of all the other epithets [such as "immortal"] attributed to it .... To *associate* soul and the Idea of Life, as happens in the final argument of the *Phaedo*, does little to resolve the difficulty; to *be* the Idea of Life, not simply to be associated with it, is essential if the soul is to be demonstrated to be immortal.[40]

That the soul shares a "kinship" with the Forms may suffice to rebut Cebes' concern about the *immediate* dissipation of the soul upon separation from the body. But at best this argument can only conclude that the soul is long-lived, not, as was its intent, that the soul is immortal.

4. The entailment argument (102a-107b)

The final sections of the *Phaedo* contain a fourth argument for the soul's immortality. It states as initial premises that the presence of the soul is necessary for a body to be alive, and that a thing will "never admit the opposite of what it introduces."[41] The soul, in other words, being defined essentially as that which gives life cannot admit the contrary principle, death. The soul, being what it is (and, as previously stated, taking for granted *that* it is), if it is at all must therefore be deathless. From this analytic conclusion Socrates goes on to extrapolate the further needed results that the deathless soul is also indestructible and imperishable.

---

[40]Robinson, *Plato's Psychology*, 48-49.

[41]Plato, *Phaedo*, 105d.

One might show this argument to be either trivial or inconsistent. On the one hand, we are not obligated to accept his definition of soul as "that which must be present in a body for it to be alive."[42] We can acknowledge the relationship while asserting the weaker claim that the soul is what all living things have. In other words, what for Plato is causal we need only accept as correlational. The distinction is whether soul is the source of life (i.e., the life-principle), or whether it is something that living things tend to have.

At issue is the *kind* of soul that Plato wishes to defend. The Formal argument depicts the soul as the rational powers capable of contemplating the Forms; such a soul, one assumes, is restricted to human beings. The soul required by the argument from Entailment, being the life-principle, is necessarily present in all living things, including plants. As their extensions are different, the two souls are not the same kinds of souls, even if Plato seems not to notice the difference.

We need not decide here which conception of the soul is the more fruitful for philosophical contemplation. The point need only be highlighted that they cannot be invoked simultaneously. The soul of the Entailment argument is too generic to contemplate the Forms; the soul of the Formal argument is too specific to be the life-principle of plants. Their dual assertion undermines both.

---

[42]Ibid., 105c.

B. Criterion Two: Kind of Soul

Having concluded the critical review of Plato's attempted proof of the immortality of the soul, we may pursue a bit further the topic introduced in the last section, the *kind* of soul at issue in the *Phaedo*. For most persons, the interest in the question of the immortality of the soul is the implied understanding that "I" will survive *by means of* the soul. To ask, "does the soul survive death?" is usually meant to mean "will *I* survive death?" The dialogue recognizes this added burden of proof: "to show that the soul exists when the man has died, *and* possesses some power and intelligence — well, that, I feel, needs a great deal of persuasive argument."[43] Even had Plato's proofs succeeded in the first goal, would he have proved the second? Neither rational power nor life-principle captures the sense of "self" which would satisfy those wondering about their postmortem survival. While rationality and life-principle are "parts" of what we are, neither is equivalent to the "I".

Plato clearly expected that his proofs of the soul were for the personal soul.[44] In every myth he creates describing the soul, that soul is exactly like a person, demonstrating the precise kind of total survival

---

[43]Ibid., 70b.

[44]Many scholars in fact believe that the equation between soul and self is an underlying theme in Plato's works. For example, Sorabji ("Soul and Self," p. 14) writes that "The self was not standardly equated in antiquity with soul. This was only one view among others, a view particularly connected with Plato."

of the personality we crave.[45] The problem, therefore, is that "Plato is prepared to accept in his myths a view of soul which stems from religious circles and is often apparently at variance with much of what he has to say about soul in the body of the dialogues."[46] "The 'true self' conception of the soul," in fact, "plays virtually no part in the arguments for its immortality."[47] Hackforth concludes that although all of Plato's proofs are designed to apply to the personal soul, in fact every one fails on this point.[48] For example,

> Plato intends the [argument from Opposition] to prove *personal* immortality, an eternal existence of individual souls eternally retaining their identity. In fact, however (and this is true of all of his arguments for immortality, with the doubtful exception of the argument from Recollection, soon to follow), it cannot be held to do so; that souls persist through the cycle as individuals souls is simply assumed: it

---

[45]Cf. the Myth of Er: Plato, *Republic*, Book X.

[46]Robinson, *Plato's Psychology*, 128.

[47]Gallop, *Phaedo*, xvi.

[48]This too is the conclusion of the medieval philosopher and theologian, William of Auvergne:

> We pass over, however, the basic arguments and proofs of Plato, because for those with intelligence they do not produce conviction about the immortality of our souls and they are common to all species of souls so that they are also applicable to the brute and to the vegetative soul. It is clear that the existence of these souls after the body and apart from the body is pointless and utterly useless.

William of Auvergne, *The Immortality of the Soul* (Milwaukee, WI: Marquette University Press, 1991), 33-34.

> is not a necessary part or corollary of the principle that 'living' comes from 'dead', or that the process of dying is balanced by the process of coming to life.[49]

The point being made is that even if we accepted the argument from Opposition, that would only prove that living comes from dead. It is not required of this cycle that the living coming out of the dead *be the same* life (soul) that existed earlier. Perhaps soul elements are recombined before rebirth, making the new life something totally new. Nothing in the Opposition argument precludes this possibility; it merely assumes that the same soul (personality, memories, etc.) survives the life-death-life cycle intact.

> And what of the argument from Recollection?
> This argument comes nearer than the previous one to being a proof of individual mortality: it is *my* soul which recollects what *I* knew before birth in the body. Yet in default of recollection of personal experience it is difficult to see how there can be that consciousness of identity preserved through a series of incarnations without which we cannot properly speak of individual immortality. As Cornford says, 'the memory implied in the doctrine of *Anamnesis* is an impersonal memory: its contents are the same in all human beings.[50]

Each offered proof, then, has its particular shortcomings relative to the demand that the soul proved to be immortal is "me": The soul of the Opposition argument does not require any personal attributes; the attributes of soul of the Recollection argument are utterly generic; the

---

[49]Hackforth, *Plato's Phaedo*, 65.

[50]Ibid., 76.

soul of the Formal argument is absolutely rational and hence not like us at all; and the soul of the argument from Entailment is a possession of all living things, not just humans.

We should not be surprised by this outcome. At the outset Socrates expresses a dichotomy which, if maintained, could never generate a soul we would recognize as being truly "us." In the *Phaedo* the desires and appetites are wholly the actions of the body, and are alien to and in competition with the rational pursuits of the soul. While the Formal argument especially depends upon this relationship, it had been introduced much earlier when Socrates is explaining why the philosopher particularly should be prepared for death: he has already at least partly overcome the corrupting influence of the body's desires. Any concept of the soul which so alienates the emotional side of human existence cannot be one which we would accept as preserving our sense of self beyond the death experience. Some of us, at least, consider our emotional side to be our better aspect, not a flaw to be minimized. Even where the emotions are not themselves virtues (as in friendships, perhaps), they are the motivators of all our actions. It is a bad psychology to argue that we should overcome our emotions; rather should we cultivate them properly.

Plato rectifies this ejection of the embodied emotions from the self when in the *Republic* he divides the soul into three parts, one of which

contains the desires and appetites. Thus what in the *Phaedo* is a competition between body and soul becomes in the *Republic* a struggle between the different levels of the soul itself.[51]

C. Criterion Three: Psychological Continuity

At the first look, Plato's soul concept seems to be highly compatible with evolutionary theory. In both the *Phaedo*[52] and the *Republic*[53] he

---

[51]Cf. Robinson, *Plato's Psychology*, 43. In a weird way, I seem to have wound up here in exactly the opposite place from where Sorabji ends. My reasoning is that the self in the *Phaedo* unnaturally separates the emotions and the reason by restricting them to the body and the soul respectively, and that this separation is alleviated for the better in the *Republic* when Plato places both emotions and reason in the soul. Sorabji ("Soul and Self," 20) has it just the other way around:

> Although Plato identifies soul and self, *Republic* Book 10 suggests that the true soul and the true self may be only that part of the soul which loves wisdom (*philosophia*) in other words, the rational part.

At the least, this conflict shows that the solution to these problems is far from obvious.

[52]"It is certainly not the souls of the righteous, but those of the wicked that are compelled to wander about such places [as ghosts], as the penalty for bad nurture in the past. And they must continue to wander until they are once more chained up in a body...; and naturally they will be chained to the type of character they have trained themselves to exhibit in their lifetime.

> What types have you in mind, Socrates?
> I mean, for example, that those who have trained themselves in gluttony, unchastity and drunkenness, instead of carefully avoiding them, will naturally join the company of donkeys or some such creatures, will they not? .... Whereas those who have set more value upon injuring and plundering and tyrannising over their fellows will join the wolves and hawks and kites.... [They] that have

allows for the transmigration of souls between animals and humans.[54] It might be thought that this is what we are looking for, a theory of the soul which does not erect impermeable barriers between "humans" and "animals."

Upon closer consideration, however, Plato seems to have ventured too far in the contrary direction: He does not place *enough* of a barrier between humans and animals.[55] Whereas for evolutionary theorists

---

> practised the common virtues of social life…will naturally find themselves in another well-conducted society resembling their old one, a society of bees, perhaps, or wasps or ants; and later on they may rejoin the human race they have left, and turn into respectable men.

Phaedo, 81d-82b.

[53]"Still other souls changed from animals into human beings, or from one kind of animal into another, with unjust people changing into wild animals and just people into tame ones." *Republic*, 620d.

[54]The kind of animal into which one is reincarnated is not random. Assuming that the soul is such as is always to be in motion, and that the proper motion for the soul is circular, a poorly trained soul will deviate from its proper circular motion. Each kind of character flaw has its own kind of deviation; the animal form is one such that can "make room" for these abnormal soul paths.

> If the circular movements of the rational part of the soul get distorted by poor use of reason we may be reincarnated as animals with a long snout, to house the distorted movements.

Sorabji, "Soul and Self," 11.

[55]He at least does not seem to go so far as Empedocles, who believed "he had once been incarnated as a bush." Sorabji, "Soul and Self," 9.

animals and humans are distinct but related, for Plato they are interchangeable, at least as far as the soul is concerned.[56] The soul of a human is just as comfortable in the body of a wasp or donkey as it is in the body of a man or a woman. How reasonable is that?

Not very, according to Thomas Nagel. Speaking in the context of the debate about animal consciousness, he argues that while such animal consciousness exists, it is beyond our ken. Using as his example the consciousness of bats, he explains that bat sonar is a sense-perception utterly unlike any of our own, and hence a mentality constructed on such input would be incomprehensible to us. "Reflections on what it is like to be a bat seems to lead us, therefore, to the conclusion that there are facts that do not consist in the truth of

---

[56]Burnyeat finds that in the *Theaetetus* Plato has leapt to the other side of this divide. While in the *Phaedo* humans and animals are of a kind, in the *Theaetetus* they are incompatible.

> [With] perception so barely characterized [in sections 184a-186e], what sense can be made of the purposive behavior of nonhuman animals who have perception only, without the coordinating capacity of interpretative judgement (186bc)? If Protagoras appeared to obliterate the difference in cognitive powers between man and beast (161c, 171e with 154a, 162e, 167b) [and as Socrates seems to suggest in the *Phaedo*], Plato has gone to the opposite extreme and made it impossible to speak of animals using their senses to get information about the environment.

Burnyeat, *The Theaetetus of Plato*, 63.

propositions expressible in human language."[57] In other words, animals and insects are not minds like ours, only housed in different bodies.[58] The different senses inherent in those different bodies generate a different kind of mind. Therefore, to the extent that rational mind translates into soul (and for the *Phaedo* that correlation is virtually complete) the soul of a donkey cannot be the same kind of soul as a human's. So long as the soul contains (or equals) the mind, transmigration is incompatible with biological reality.

Can the thesis of soul transmigration be salvaged if we severe the connection between soul and mind? Perhaps, but it is difficult to see

---

[57]Thomas Nagel, *Mortal Questions* (Cambridge: Cambridge University Press, 1979), 171. We should note that not everyone agrees with Nagel's pessimism on this point. Daniel Dennett, for one, disputes this negative conclusion. He feels that Nagel "didn't so much argue for this conclusion as assume it and then discuss its implications." Daniel Dennett, *Brainchildren* (Cambridge, MA: MIT Press, 1998), 337. I, however, find Nagel's position to be eminently reasonable.

[58]A possible retort: Nagel's discussion is about bats because their sensory apparatus is radically different from our own. Does it apply to donkeys? Maybe what it's like to be a donkey is just what it's like to be an extremely dumb person. If true, that would perhaps blunt the force of the criticism being offered here.

But Nagel offers no suggestion that I can find that he is saying that only some kinds of animal consciousness are beyond our ken. And certainly, according to Daniel Dennett, the practical effect of the article has been to shut down all research into animal consciousness, not just into certain kinds. Daniel Dennett, *Brainchildren*, 337. So while the argument could be made that only some kinds of animal consciousness (those whose sensory apparatus is utter alien to our own) is beyond our reach, others not, this is not the argument I believe Nagel is actually making here.

how that could be done without transgressing the other of our requirements, that the soul in play be "self-preserving." The self minimally includes our memories, and these are intimately connected (if not actually lodged in) the mind. To lose the mind as an aspect of soul, then, erases the memories which define our "selves." Consequently, in the process of the soul's migration it is shorn of any impress of having been "my" soul, and amounts to little more than the recycling of generic soul-stuff. This kind of soul-survival is frankly uninteresting, and is emotionally indistinguishable from the kind of survival typical of a transplanted organ.

Again, the question is not whether "part" of me survives, but whether "I" survive. If the surviving soul is deemed capable of transmigration, it is a soul theory in conflict with biological theory; if the soul is altered in such a way as to be compatible with biological theory, then it is no longer a soul which preserves the self.

D. Conclusions

As the Part I of this thesis argued, soul-theories are properly critiqued by three standards: internal self-consistency, reflection of the common intuitions about the self, and compatibility with evolutionary theory. Plato's model of the soul as defended in the *Phaedo* fails on all three prongs.

As we saw, each of the four attempts to prove the soul's immortality conflicts with other aspects of the wider Platonic philosophy. Worse, some of the offered proofs conflict with other of the proofs in this same dialogue. For example, the Formal argument and the Entailment argument cannot both be valid if "soul" has the same meaning for both proofs. Whereas the Formal argument requires that the soul be such as to be able to contemplate the Forms, the Entailment argument works with a generic life-principle present in all living things. This life-principle is both more extensive and less complex than the cognitive powers necessary for contemplating the Forms. The soul as conceived for one argument would fatally confound the arguments of the second.

We also concluded that even were the proofs convincing, we wouldn't much care. Each proof required a slightly different characterization of the soul, but none of them are equivalent to our fuller sense of "self."[59] The recurring shortcoming on this issue was the strict bifurcation of the soul/body being overlaid on an equally strict bifurcation of mind/emotions. Since most persons would accept emotions into the sense of self they wish preserved after death, and as the basic concept of soul in the *Phaedo* does not allow for this inclusion,

---

[59]Robinson (*Plato's Psychology*, 30) counts five different concepts of the soul in this dialogue: counter-person, intellectual principle, life-principle, formal property, and ectoplasmic. This book's Coda outlines all the sundry formulations of the soul discernible through the entire Platonic corpus.

then no such soul theory would satisfy the needs which initiated the quest.

Finally, while Plato's allowance of the transmigration of souls between biological species seems to be compatible with evolutionary theory, in fact it errs too far in the opposite direction. It does not make *enough* distinction between animals and humans, as contrasted with other theories we shall examine, which overemphasize this distinction.

CHAPTER III

ARISTOTLE'S *DE ANIMA* (c. 325 B.C.E.)

Especially as compared to Plato, where the danger is that the pleasant language and vivid dramatic imagery can distract from the deeper philosophical arguments, Aristotle's texts can be dense and excruciatingly technical: "reading Aristotle, as the poet Thomas Gray put it, is like eating dried hay."[1] The effort, if intimidating, is however almost always rewarded.

The *De Anima* ("On the Soul") is part of the exhaustive coverage of Aristotle's writings. The *Physics* considers the natural, physical world, and the *Metaphysics* considers being *qua* being; the *De Anima* is the work which is in many ways conceptually intermediate between these two, devoting its pages to the study of those physical things which are alive.

As we shall see, "the conception of the soul as the entelechy of the living body is…the central conception of the *De Anima*."[2] This thesis claims that the soul is what makes the living body a living body; it is the body's first actuality. Under these conditions there can be no self-

---

[1]Jonathan Barnes, "Life and Work," in *The Cambridge Companion to Aristotle*, edited by Jonathan Barnes (New York: Cambridge University Press, 1995), 12.

[2]David Ross, *Aristotle: De Anima* (Oxford: Oxford University Press, 1961), 10.

preserving soul. However, Aristotle also claims that there is indeed a part of the soul that survives death, what he calls the "active mind." This simultaneous assertion compels us to adjudge the argument of the *De Anima* internally inconsistent.

Although Aristotle earns high marks for his overall endorsement of the principle of psychological continuity, the conflicting soul theories both run up against the second criterion of the kind of soul being defended. If the soul is truly the entelechy of the living body, then there is no soul at all, and hence no self that is preserved. And even if this definition of the soul is not absolute, and some part of the soul (i.e., the active mind) does survive death, this soul is wholly rational, and fails to capture our intuitive notion of the self.

A. Criterion One: Theory Consistency

Aristotle's *De Anima* lies midway between the proto-Cartesian dualism found in Plato's *Phaedo*, and the materialism displayed by his own contemporaries, the atomists.[3] The soul for Aristotle, as with the materialists, is inseparable from the body (with one problematic exception we will consider below), but like the dualists it is not reducible to the body.

---

[3]Cf. M.F. Burnyeat, "Is an Aristotelian Philosophy of Mind Still Credible (A Draft)?," in *Essays on Aristotle's De Anima*, edited by Martha Nussbaum and Amélie Oksenberg Rorty (New York: Oxford University Press, 1992), 16.

He arrived at this third approach to the soul because he asks a new question. Others asked, What is the soul (mind, consciousness)?, and responded in materialistic or dualistic ways. The point is that the question already separates out the soul from the body, leaving it little wonder that the answer is framed in similar terms which highlight the soul's distinctiveness.

Beginning with a severable soul requires one to confront the problem of somehow joining the soul with the body. Aristotle labels prior attempts on this problem an "absurdity":

> The view we have just been examining, in company with most theories about the soul, involves the following absurdity: they all join the soul to a body, or place it in a body, without adding any specification of the reason of their union, or of the bodily conditions required for it.[4]

We are particularly attuned to this difficulty from our acquaintance with later Cartesianism.[5]

---

[4]Aristotle, "On the Soul," translated by J.A. Smith, in *The Basic Works of Aristotle*, edited by Richard McKeon (New York: Random House, 1941), 407b13-17. All quoted passages from this treatise come from this translation.

[5]E.g., "Both Aristotle and the scholastic Aristotelians believe that the soul is that in virtue of which a living body is alive. Descartes, on the other hand, assumes that bodies, whether alive or not, form part of the physical world and can be explained in its terms without recourse to such entities as souls." Michael Frede, "On Aristotle's Conception of the Soul, in *Essays on Aristotle's De Anima*, edited by Martha Nussbaum and Amélie Oksenberg Rorty (New York: Oxford University Press, 1992), 94. For Descartes, then, the problem will be how a perfectly immaterial soul can interact with a material body.

Aristotle, however, asks something else: "At the core of his enterprise is the concern to characterize the explanatory principles that make living things different from non-living one."[6] Consequently "the 'soul' is construed as that which enables a natural body to be a living thing."[7] Whereas earlier the body had been characterized as an independently existing material thing to which a soul is somehow appended, now the soul is that which makes a body a 'body' at all.

The soul (*psuche*) is "the first grade of actuality (*entelecheia*) of a natural body having life potentially in it."[8]

> *Psuche* is — as common speech has it — the life and soul of an organism, engaged in its natural activities. An organism does not have life as one of its attributes, along with its size and shape. Rather, the life and soul of a certain kind of body consists in its being active in a certain way, engaged in those activities that constitute its being the sort of thing it is. Life is not a presupposition of activity; rather, to be alive is to be actively (endogenously) engaged in those activities which constitute one's nature.[9]

---

[6]Alan Code and Julius Moravesik, "Explaining Various forms of Living," in *Essays on Aristotle's De Anima*, edited by Martha Nussbaum and Amélie Oksenberg Rorty (New York: Oxford University Press, 1992), 132.

[7]Ibid.

[8]Aristotle, "On the Soul," 412a28.

[9]Amélie Oksenberg Rorty, "*De Anima*: Its Agenda and Its Recent Interpreters," in *Essays on Aristotle's De Anima*, edited by Martha Nussbaum and Amélie Oksenberg Rorty (New York: Oxford University Press, 1992), 9-10.

This change in presupposition about the soul permits Aristotle to avoid the absurdity for which he faults his predecessors. Because the soul is "the first grade of actuality of a natural organized body,"

> we can wholly dismiss as unnecessary the question whether the soul and the body are one: it is as meaningless as to ask whether the wax and the shape given to it by the stamp are one, or generally the matter of a thing and that of which it is the matter.[10]

The immediate point to stress is that due to this hylomorphism, "it indubitably follows that the soul is inseparable from its body."[11]

> [The] body cannot be the actuality of the soul; it is the soul which is the actuality of a certain kind of body. Hence the rightness of the view that *the soul cannot be without a body*, while it cannot *be* a body; it is not a body but something relative to a body.[12]

By this definition, then, there can be no self-preserving soul which survives death. The soul and the body are conceptually distinguishable (so again, Aristotle is not a materialist), but they are also as inseparable as is the wax and its shape (so he is not a dualist). To have a soul is to be engaging in the kinds of acts which besuit besouled matter. An eye that cannot see is not an eye;[13] a body without a soul is not an animal[14]

---

[10]Aristotle, "On the Soul," 412b5-8.

[11]Ibid., 413a4.

[12]Ibid., 414a16-20 (emphasis added).

[13]Ibid., 412b18-23. Cf. Aristotle, "De Partibus Animalium," translated by William Ogle, in *The Basic Works of Aristotle*, edited by Richard McKeon (New York: Random House, 1941), 641a3: "no part of a

or even a body. The soul is what actualizes the body, and from that perspective it makes no sense to speak of a bodiless soul.

The bulk of the *De Anima* is devoted to the task of drawing out the implications of this understanding of the soul as the body's first actuality.[15] Because being alive and having a soul is nothing but the ability to engage in the actions characteristic of the kind of thing that it is, Aristotle spends much of the text considering how one does the things one does (e.g., how does one see, touch, feel, think). Additionally, by this definition souls are not restricted to humans, and so he considers also animal and plant souls, seeking to isolate the faculties which make them

---

dead body, such I mean as its eye or its hand, is really an eye or a hand;" and "On Interpretation," translated by E.M. Edghill, in *The Basic Works of Aristotle*, edited by Richard McKeon (New York: Random House, 1941), 21a22: "it is not right to call a dead man a man."

[14]Aristotle, "On the Soul," 413a3.

[15]Perhaps the best brief statement of Aristotle's position on the soul is the following from Ross, *Aristotle: De Anima*, 216:

> Therefore they judge well who think the soul neither exists apart from body nor is body; it is not a body, but something belonging to a body; it therefore exists in a body, and in a body of a particular kind. People used to fit soul into body without defining the kind of body required, though it does not look as if any chance body could receive any chance capability. We do so no longer, and that is reasonable; for it is only in appropriate matter that an actuality can be lodged. Evidently, then, soul is an actuality, the formulable essence of that which can assume such and such a character.

the kinds of things that they are, as empowered by their appropriate souls.

These implications have generated a voluminous literature which it is beyond the ability of the present section to examine. It can be said, however, that while weaknesses identified here may be many (e.g., the model builds upon an erroneous understanding of the physics of the senses[16]) these are not the kinds of flaws we are identifying as crippling the theory of the soul. They are too far removed; our task, recall is to determine whether the soul theory has any *a priori* validity, whether or not it is obviously inconsistent on its own terms.

To answer that question, we have to consider Aristotle's claim that "If there is any way of acting or being acted upon proper to the soul, soul will be capable of independent existence; if there is none, its separate existence is impossible."[17] He concludes, as he must for his thesis, that "there seems to be no case in which the soul can act or be acted upon without involving the body,"[18] that "it is obvious that the affections of the soul are enmattered formulable essences."[19]

---

[16]Burnyeat, "An Aristotelian Philosophy of Mind," 16.

[17]Ibid., 403a10.

[18]Ibid., 403a5.

[19]Ibid., 403a24.

With one exception: "Thinking seems the most probable exception,"[20] meaning that, "not having realized the importance of the brain, [Aristotle] assigns no bodily seat to reason."[21] But recall that "if there is *any* way of acting … proper to the soul [i.e., independent of the body], the soul will be capable of independent existence." If thinking is a way of acting proper to the soul, then the soul has an existence independent of the body, contradicting Aristotle's claim that the soul is the first actuality of the body.

Aristotle's distinction between the passive and the active intellects is relevant here. Here we find him again hedging on the independence of the soul. He first analogizes thinking to perceiving, the discussion of which has consumed much of the preceding text.

> If thinking is like perceiving, it must be either a process in which the soul is acted upon by what is capable of being thought, or a process different from by analogous to that. The thinking part of the soul must therefore be, while impassable, capable of receiving the form of an object; that is, must be potentially identical in character with it subject without being the object. Mind must be related to what is thinkable, as sense is to what is sensible.[22]

Thinking is, by this description, a "passive affection,"[23]

---

[20]Ibid., 403a7.

[21]Ross, *Aristotle: De Anima*, 42.

[22]Aristotle, "On the Soul," 429a13-17.

[23]Ibid., 429b23.

reacting to the forms presented to it.

The passive intellect is to be contrasted with the active intellect.

> And in fact mind as we have described it is what it is by virtue of becoming all things [the passive intellect], while there is another which is what it is by virtue of making all things [the active intellect]: this is a sort of positive state like light; for in a sense light makes potential colours into actual colours.[24]

The active intellect [*nous poietikos*] is like a light: turn it on and suddenly you can see colors; the actual seeing of the colors is what corresponds to the passive intellect. By this analogy it is the passive intellect which does the thinking, but it is the active intellect which enables the passive intellect to do this thinking. The [active intellect] is said to exist by making all things, and the [passive intellect] by becoming all things."[25]

The active intellect is not only functionally distinct from the passive intellect, but it also differs in terms of its long-term fate:

> When [the active] mind is set free from its present conditions it appears as just what it is and nothing more: this alone is immortal and eternal (we do not, however, remember its former activity because, while mind in this sense is impassible, mind as passive is destructible), and without it nothing thinks.[26]

---

[24]Ibid., 430a14-17.

[25]Ross, *Aristotle: De Anima*, 45.

[26]Aristotle, "On the Soul," 430a22-25. David Ross (*Aristotle: De Anima*, 194) explains the loss of memories this way:

> Thus thought decays through the destruction of something within us; it does not itself suffer. Reasoning, love, hate are affections not of it but of that which possess it, *qua*

So something in the soul is, after all, independent of the body, and immortal. This something is not, as was first intimated, thinking, but instead the active principle which makes thinking possible.

The question for us is whether it even makes sense to define the soul as Aristotle has, as the first actuality of the living body, and then to attach onto this kind of soul an immortal essence which is *not* a body's first actuality. Scholars working in this area (e.g., Ross[27]) may concede the conceptual difficulty in reconciling these two claims, but ultimately find a way to do so.

Surely it cannot be the case that we have two souls. Nor, soul being what it has been defined to be, there can there be any immortal remainder. So how should the active intellect described in Chapter 3.5 to be reconciled with the rest of the treatise, if it all?

On the one side we have the refreshing honesty of K.V. Wilkes, who admits to being "puzzled and embarrassed by *DA* 3.5." "I cannot understand this chapter, and none of the secondary literature has so far

---

possessing it. That is why, when this possessor perishes, we do not remember or love; these activities belonged not to reason but to the composite entity which has perished.

[27]For Ross' attempted reconciliation of the active and passive minds, see his *Aristotle: De Anima*, 46-47.

helped me to do so;" indeed, he wishes Aristotle "had never written this chapter."[28]

There is certainly a sense in which 3.5 is superfluous to the arguments of the *De Anima*. The section is short, and adds little or nothing to the problems which are the treatise's main concern. Perhaps 3.5 is an anomaly which is more properly ignored than analyzed. Jonathan Barnes warns us that no treatise can be interpreted with "the assumption that it is an Aristotelian unity."[29] Maybe this problematic section is due to inappropriate editorial insertion. David Ross suggests that not only is the text carelessly written, but that "it is possible that parts have disappeared which would have made the chapter more intelligible."[30] So perhaps there is no genuine philosophical problem

---

[28]K.V. Wilkes, "Psuche versus the Mind," in *Essays on Aristotle's De Anima*, edited by Martha Nussbaum and Amélie Oksenberg Rorty (New York: Oxford University Press, 1992), 111, 125.

[29]Barnes, "Life and Work," 11. Whereas the problem in Plato was how to relate his thinking between texts, the major problem here relates to the nature of the individual text itself:

> You cannot read Aristotle in the way in which you might read Plato or Descartes or Kant: when you pick up the *Metaphysics* or the *Nicomachean Ethics*, you are not picking up a finished philosophical text, comparable to the *Theaetetus* or the *Meditations* or the *Critique of Pure Reason*. It is proper to assume that you are picking up a set of papers united by a later editor; and it is proper to assume that you are reading a compilation of Aristotle's working drafts. (p. 15)

[30]Ross, *Aristotle: De Anima*, 296.

here, but only an apparent one due to textual corruption or editorial errors.

While this solution would preserve the consistency of the *De Anima*, before we endorse it we should first consider another possibility. Maybe the section's primary function is not to further the thesis of the *De Anima* but instead to provide a bridge between this work and others of the Aristotelian corpus, particularly the *Metaphysics*. If this is true, then 3.5 is not a textual anomaly, but a claim necessitated by the wider Aristotelian philosophy. If 3.5 is performing this important work, it then presents a genuine problem. If this section cannot be ignored, it renders inconsistent the assertion of hylomorphism. What grounds have we to prefer this second interpretation?

The passage in 3.5 which attracts our attention in this context is the following:

> Actual knowledge is identical with its object: in the individual, potential knowledge is in time prior to actual knowledge, but in the universe as a whole it is not prior even in time. Mind is not at one time knowing and at another not.[31]

Paralleling the active/passive distinction, then, is the actual/potential.

---

[31]Aristotle, "On the Soul," 430a20-22. It was suggested above that one of the facts contributing to the position that 3.5 is a negligible distraction is its sheer brevity. It is therefore worth noting that with the present extract we have here reproduced 3.5 in almost its entirety; that's how short it is.

The passive intellect depends upon the omnipotentiality of the mind. That is to say, because the mind can think all things, it is nothing in itself, again reasoning from analogy with the senses: "Thus that in the soul which is called mind is, before it thinks, not actually any real thing."[32]

To "think" is to move from potential mind to actual mind. However, as explained in the *Metaphysics*, "from the potentially existing the actually existing is always produced by an actually existing thing...; *there is always a first mover, and the mover already exists actually*.[33] The potential of the mind to think, therefore, presupposes a prior actuality: "Obviously, then, actuality is prior both to potency and to every principle of change."[34]

The passive intellect, then, is a potency which presupposes a prior actuality, which Aristotle terms the *nous poietikos*, or active intellect. The actuality of the active intellect in the sense of being a first mover

---

[32]Ibid., 429a23.

[33]Aristotle, "Metaphysics," translated by W.D. Ross, in *The Basic Works of Aristotle*, edited by Richard McKeon (New York: Random House, 1941), 1049b23-25 (emphasis added).

This position is repeated in the *De Anima*: "Actual knowledge is identical with its object: potential knowledge in the individual is in time prior to actual knowledge but in the universe it has no priority even in time; for all things that come into being arise from what actually is." Aristotle, "On the Soul," 431a1-3.

[34]Aristotle, "Metaphysics," 1051a2.

must by definition exist always, because what is actual cannot be potential. In this sense, then, the actuality of the active mind must be eternal: "eternal things are prior in substance to perishable things, and no eternal thing exists potentially."[35]

This is not to say (as a third possibility) that the principle of the active mind cannot co-exist with the soul; only that it is itself definitionally not a soul, or part of a soul, but something else altogether. We need not worry here just what that something else might be. Michael Frede describes this possibility: The activity of the active intellect

> clearly is not the exercise of a natural ability of a body, and thus its thought clearly cannot be explained in the way, I suggested earlier, Aristotle wants to explain the mental functions of human beings. But we can acknowledge this, without having to draw the conclusion that Aristotle's doctrine of the active intellect is incompatible with the view that the soul is just the form of a natural body. For it is open to us to assume, following in this a long tradition of interpreters, that this active intellect is not a human intellect, that it is not an integral part of the human soul. Thus there also is no need why Aristotle, given his view of the human soul as the form of a body, should be able to explain the activity of the active intellect along the lines he tries to explain the mental functions of human beings.[36]

The argument then is this: Aristotle's conception of the soul is designed to explain *natural* capacities; the active intellect is a *supernatural* capacity, and hence the failure of his model of entelechy to include it is

---

[35]Ibid., 1050b6-8.

[36]Frede, "On Aristotle's Conception of the Soul," 105.

not a fault of the theory. Inconsistency is avoided by redefining the problem. Such resort to a psychological *deus ex machina* would, however, be an uncharacteristic lapse for Aristotle.

Still, for our purposes, either of two results apply. If the active mind is intended to be a soul, this assertion is incompatible with Aristotle's definition of "soul," and hence the soul-theory articulated in the *De Anima* is inconsistent; or, if the active mind is something other than a soul, and if Aristotle's calling it by that name is just loose language, the theory is a consistent argument *against* the existence of a self-preserving soul.[37] On no reading of the *De Anima* can one extract support for the post-mortem survival of the personality.

B. Criterion Two: Kind of Soul

From our perspective it perhaps matters little which version of Aristotle's soul theory should prevail, since neither offers a model of a

---

[37]Given the hallowed status of Aristotle, it merits mention that the conclusion that his characterization of the soul is fatally inconsistent is not unique. Robin Smith, for instance, finds that "Aristotle's system [of the modal syllogistic] is incoherent and that no amount of tinkering can rescue it." "Logic," in *The Cambridge Companion to Aristotle*, edited by Jonathan Barnes (New York: Cambridge University Press, 1995), 45.

C.C.W. Taylor ("Politics," in *The Cambridge Companion to Aristotle*, edited by Jonathan Barnes [New York: Cambridge University Press, 1995]) identifies similar inconsistencies in the *Politics*: "The thesis that the *polis* stands to the individual as whole to part is ... an aberration on Aristotle's part; it commits him to denying two central theses of his ethico-political system" (p. 241).

postmortem self-preserving soul. They differ only on which variable is emphasized to the complete exclusion of the others.

The theory as Aristotle intended — soul as first actuality, hylomorphic — completely captures the need for a soul which adequately contains the concept of "self." By this first reading, where the soul is, there is everything that we take ourselves to be.

Unfortunately, this total picture is achieved by arguing that the soul/body difference is sustainable in theory only, and is not reflected in the world. That is to say, while we can speak of the soul as distinct from the body, they do not actually occur in this way. The result is that we have maximized the possibility for a soul which contains our self, at the cost of making that soul perishable and thus essentially unable to be preserving of that self after death.

The idea of the active intellect, as we saw, reintroduces the dualistic idea that some part of the soul does survive death after all. So while we now have a soul that preserves, it is not the self which is preserved. If the active intellect is indeed like light,

> this makes it hard to consider the active intellect as being itself a form *of thought*; rather, it seems to be what makes thinking possible. If indeed this is what Aristotle says to be 'immortal and eternal', it is difficult to get personally engaged in, or excited by, its immortality: it would be nothing like *me* that survives bodily death.[38]

---

[38]Wilkes, "Psuche versus the Mind," 125.

This version of immortality "is more like that of Heraclitus than that of contemporary Christianity: we (you, I, Heraclitus) do not survive as ourselves, but something else does: the Heraclitean fire, or the Aristotelian light of the intellect."[39] So even if something does survive after all, it is plainly not the self. So regardless of which interpretation of the soul we decide is the properly Aristotelian, neither satisfies our search for a post-mortem self-preserving soul.

C. Criterion Three: Psychological Continuity

The *De Anima* text is ambivalent on the question of psychological continuity. How one classifies Aristotle's positions depends on where one puts the interpretive emphasis.

Favoring the principle of psychological continuity is the fact that while plants, animals, and humans have different kinds of souls (the nutritive/reproductive, the sensitive, and the mind, respectively), the soul of each higher creature includes the powers of those below it, although these powers are differently instantiated. For example, while touch is the only sense that every animal must possess to live, the mechanics of touching varies from species to species.

The proper understanding of sense perception in animals will thus be a complete explanation of sense perception in humans. On the other hand, an understanding of sense perception in animals is a complete

---

[39]Ibid., 126.

understanding of what it is to be an animal, but not of what it is to be human. We, so Aristotle asserts, are animals in one sense, but we are not merely a kind of animal, coordinate with horses and cows. We are more than animals owing to our possession of a kind of soul different from any mere animal, one which lets us think. Here then seems to be a difference in kind which denies the principle of psychological continuity.

Other claims outside the *De Anima* tip the balance in favor of support for psychological continuity. On the one hand we read the unqualified statement that "other animals than man have the power of locomotion, but in none but him is there intellect."[40] By contrast there is the following extended passage from the *History of Animals*, which remarkably presages the position advocated by Darwin in *The Descent of Man*:

> In the great majority of animals there are traces of psychical qualities or attitudes, which qualities are more markedly differentiated in the case of human beings. For just as we pointed out resemblances in the physical organs, so in a number of animals we observe gentleness or fierceness, mildness or cross temper, courage or timidity, fear or confidence, high spirit or low cunning, and, with regard to intelligence, something equivalent to sagacity. Some of these qualities in man, as compared with the corresponding qualities in animals, differ only quantitatively: that is to say, a man has more or less of this quality, and an animal has more or less of some other; other qualities in man are represented by analogous and not identical qualities: for instance, just as in man we find knowledge, wisdom, and sagacity, so in certain animals there exists

[40]Aristotle, "De Partibus Animalium," 641b8.

> some other natural potentiality akin to these. The truth of this statement will be the more clearly apprehended if we have regard to the phenomena of childhood: for in children may be observed the traces and seeds of what will one day be settled psychological habits, though psychologically a child hardly differs for the time being from an animal; so that one is quite justified in saying that, as regards man and animal, certain psychical qualities are identical with one another, whilst others resemble, and others are analogous to, each other.[41]

Aristotle here identifies no psychical quality or attitude which is the prerogative of man alone. Every thing that man has or does, animals likewise can claim either directly, in lesser or greater degree, or by analogy.

This includes the singular ability regarded in the *De Anima* as being the special function of the human soul, the power to think. Animals possess "something equivalent to sagacity," "some other natural potentiality akin to" man's knowledge and wisdom.

The principle of psychological continuity is further supported by Aristotle's equation of the human child with animals. The adult human, by this model, emerges out of an earlier state indistinguishable from animals, absolutely precluding any argument for a difference in kind between the two.

---

[41]Aristotle, "History of Animals," translated by D'Arcy Wentworth Thompson, in *The Basic Works of Aristotle*, edited by Richard McKeon (New York: Random House, 1941), 588a18-b3.

Finally, Aristotle concedes that due to the imperceptible gradations between forms, categorical distinctions are difficult if not impossible to maintain: "Nature proceeds little by little from things lifeless to animal life in such a way that it is impossible to determine the exact line of demarcation, nor on which side thereof an intermediate form should lie."[42] We can justly argue that a similar difficult arises on the other half of the biological scale, between animals and humans, again precluding any strict line of demarcation.

We should conclude, then, that despite the apparent ambiguity on the issue within the *De Anima*, Aristotle intends no contradiction of the principle of psychological continuity.

D. Conclusions

Aristotle's treatment of the soul fails on two of the three criteria. First, the *De Anima* contains equally strong arguments for a soul which survives death, and one which does not. While the thrust of the text is directed toward the latter, the former is a necessary addendum if this argument is to be consistent with his wider philosophical system. While multiple souls per se do not present a problem by our criterion, they do for Aristotle in that he defines "soul" in such a way that it is necessarily mortal. The necessary assertion of both kinds of soul, then, is logical inconsistency.

---

[42]Ibid., 588b4-6.

Whatever of the Aristotelian soul might survive, in no way is it anything at all like the "self" we wish preserved. At best it is only the rational part of our mind; at worst, the inspecific power which enables us to think. In either case it is not "us" in the customary sense.

And third, whatever Aristotle's conception of the soul, his understanding of the relationship between animals and humans incorporates the principle of psychological continuity. Although this conclusion is supportable by statements within the *De Anima*, so too are others to the contrary. The certainty for our result here therefore rests largely upon his clearer assertions in other texts.

CHAPTER IV

WILLIAM OF AUVERGNE'S *THE IMMORTALITY OF THE SOUL* (c. 1128-35)

William of Auvergne, Bishop of Paris, "represents the first generation of Paris masters to make a wide use of Aristotelian, Islamic, and Jewish thought... William was the first great master of this new age."[1] The Aristotelian background is particularly crucial in deciphering William's arguments; for this reason this section will repeatedly indicate the probably origin of specific ideas from Aristotle's various texts.

Only recently has *The Immortality of the Soul* been recognized as being from his pen. The paradigm version of the text had been traditionally attributed to Domenicus Gundissalinus; the copy assigned to William was understood as a massive feat of plagiarism.[2] According to Roland Teske the most convincing case for William's authorship was made only as recently as 1968.[3]

---

[1]David Knowles, "William of Auvergne," in *The Encyclopedia of Philosophy*, edited by Paul Edwards (New York: Macmillan, 1967), Vol. 8, 303.

[2]Ernest A. Moody, "William of Auvergne and His Treatise De Anima," in *Studies in Medieval Philosophy, Science, and Logic* (Berkeley, CA: University of California Press, 1975), 8.

[3]William of Auvergne, *The Immortality of the Soul*, translated by Roland J. Teske (Milwaukee, Wisconsin: Marquette University Press, 1991), 4.

Ernest Moody counts more than twenty treatises, not including the *De Immortalitate Animae*, and several hundred sermons which "are generally recognized as authentic works of William of Auvergne."[4] Of particular relevance is his *De Anima*, which devotes thirty pages to the problem of the immortality of the soul. The *De Anima* potentially provides a useful point of comparison with the arguments of the *De Immortalitate Animae*; arguments which seem cursory or poorly developed in the latter might be better appreciated in the fuller setting of the complete *De Anima*. Unfortunately the full text of the *De Anima* remains untranslated.[5] For this reason that we will restrict our consideration solely to *The Immortality of the Soul*.

The outcome of this analysis will be that the philosophical arguments are uniformly weak and underdeveloped, but almost never fatally flawed. Also, the manner in which William depicts the relationship between humans and animals is sketchy, but contains no suggestions which are not evolutionarily compatible. The major failing of William's effort, given our stated standards, is that the kind of soul defended contains too few of the qualities we normally deem adequate for a properly constituted "self."

---

[4]Moody, "William of Auvergne," 6.

[5]Significant passages are translated, however, by Moody in his article, which, while written in the 1930s as a philosophy Master's thesis, might well remain the principle source in English on that medieval text.

A. Criterion One: Theory Consistency

William of Auvergne holds that there are five ways by which human errors can be corrected: (1) sensation and experience; (2) punishment; (3) philosophy; (4) authority; and (5) revelation.[6] The "error which concerns the natural mortality of rational souls" especially warrants these correctives. William, however, limits his arguments to only the third, or philosophical method of correcting errors. He eschews invocation of religious authorities, rightly feeling that these sources would prove ineffective among his intended audience, erroneous philosophers.

The structure of *The Immortality of the Soul* is compendious. He sketches an argument in a few sentences, perhaps addressing one or two objections, quickly moving onto the next. These arguments fall into two major categories. The first contains four arguments based upon the nature of God.

It is impossible to evaluate critically the defense of an immortal soul which depends upon alleged personality traits of a presumed Supreme Being.[7] Even if successful, William has at best moved the locus

---

[6]William, *The Immortality of the Soul*, 23. All quoted passages from this treatise come from this translation. The translator's "An Outline of the Work" (pp. 8-13) was especially helpful in dissecting this piece, and is influential even where not explicitly credited.

[7]The Basingers described "the classical Christian God [as] a being who in addition to being an invisible, rational object of worship is also

of the needed rational foundation we are searching for from the question of the soul to the question of a soul-entailing God, hardly an improvement.[8] The tactic would be more relevant to our concerns had he buttressed these claims with a separate demonstration of the existence of any Supreme Being, much less one with the traits he assigns.

Given the theological assumptions characterizing this first category, we can ignore it here and move directly to the second category of arguments, the "root principles" based upon the nature and activities of the soul. Although Teske numbers these as six, only five of them are

---

considered the omnipotent, omniscient, perfectly good creator of all." David Basinger and Randall Basinger, *Philosophy and Miracle: The Contemporary Debate* (Lewiston, NY: Edwin Mellen Press, 1986), 4.

Although the argument cannot be developed here, it is arguable that the traditional attributes of the Christian God are precisely those which will entail the immortality of the soul. Thus, if God is just, and perfect justice is not available in this world, then there must be an after-life where perfect justice can be meted. These are the kinds of proofs William offers.

The interesting conclusion here would be that the very idea of God is subordinate to the idea of personal immortality. God is not the end of religious thought, but the means by which one is assured of a post-mortem survival. Failing this, it would be extremely difficult to explain why the Christian God has been assigned just the qualities he has, and no others.

[8]These are two distinct, independent problems. The existence of a God does not entail the existence of immortal self-preserving souls (unless, as suggested in footnote 7, you have also proved that the God is of a very specific kind), and the demonstration of the reality of souls is not a proof for the existence of a god of any kind.

sufficiently distinct to warrant separate consideration.[9] The five principles are those based upon Dependence, Incorruptibility, Natural Desires, Destructibility, and Contrast with the Senses.

1. The argument from dependence

William asserts as a given that "If the activity of a substance does not depend on the body, its essence does not depend upon the body."[10] His task is only to demonstrate that the human soul meets this criterion of its activities not depending on the body.

---

[9]The fourth argument according to the structure of the text is an appeal to the fact that "the life of the intellective power is not dependent upon the body." This thesis is indistinguishable from the claim in the first root principle, and has many overlapping elements. Both, for instance, cite dreams and raptures as evidence of the independence of the intellective capacity from the body. This later section does include elements not incorporated into the earlier argument (e.g., the implications from brain anatomy and injuries thereto). But nothing there responds to the identified weakness of the first root principle. So the fourth root principle offers additional reasons to support the independence of the rational soul from the body, but no new reasons to set aside the problem raised by the impede/occupy distinction. Consequently these sections are not given separate consideration here.

[10]William, *The Immortality of the Soul*, 27. Teske identifies the ultimate source of this first "root principle" as Aristotle's *De Anima* I, 1, 403a10-11: "If there is any way of acting or being acted upon proper to the soul, soul will be capable of separate existence; if there is none, its separate existence is impossible." Aristotle, "On the Soul" [*De Anima*], translated by J. A. Smith, in *The Basic Works of Aristotle*, edited by Richard McKeon (New York: Random House, 1941), 537. All quoted passages from this treatise come from this translation.

He identifies at least three justifications for this conclusion: *(a) the intellect's proper activity is not impeded by the body;*[11] *(b) the strengthening of the intellect should follow from the strengthening of the body and vice-versa, but the opposite is the case;*[12] *and (c) in its complete separation from the body, [the intellective power's] activity of the intellective power is invigorated and strengthened as indicated in ecstasy and rapture.*[13]

The considered objection is whether "the intellective power is impeded and weakened when the body is impeded and weakened, as in those who are ill, for example, in those who are delirious, out of their minds, melancholy and mentally alienated."[14] In response he writes that "to impede or injure is not the same as to occupy," and also that these emotions (like dreams) "do not harm the essence of the intellective power, but impede its activity by occupying it."[15]

---

[11]William, *The Immortality of the Soul*, 28.

[12]Id. Teske identifies the ultimate source of this point of justification as Aristotle's *De Anima*, I, 4, 408b18-25: "if [the soul] could be destroyed at all, it would be under the blunting influence of old age."

Teske includes a fourth point of justification, (d): "*unlike mortal things, the intellective power is not weakened by age*." Although a slightly different claim than that in (b), it is unclear that it is genuinely independent. If (b) is valid, then (d) must be as well; but (d) could be valid and (b) not. Therefore, for our purposes, it is necessary only to consider (b).

[13]William, *The Immortality of the Soul*, 30-31.

[14]Ibid., 29.

[15]Ibid.

This response is confused. William first distinguishes "occupying" as an *alternative* to "impeding" (i.e., "to impede or injure is not the same as to occupy"), yet later speaks of occupying as a *species* of impeding (emotions impede by occupying, not injuring). In other words, in the first case an occupied mind is neither impeded nor injured, while in the second to occupy is to impede but not to harm or injure. Only the first reading is compatible with the explicit wording of the first justification, which requires there be no species of impediment at all (regardless of whether it merely occupies or actually harms). Similarly challenged is the spirit of the root principle itself, which asserts the independence of the soul from the body.

The conclusions we draw about this first root principle are critical, because William will try later to justify others of his principles by appeal to this first one. If it is flawed, so too immediately are any others dependent upon it.

The weakness lies not in William's arguments supporting his principle, but in his attempt to rebut the objection. At best, his response is muddled and confused; at worst, it is inconsistent, and leaves the objection standing. Our project is not to worry whether the argument could be "corrected" somehow to resolve the difficulty; our only consideration is with the actual text before us. On that basis, we can cautiously conclude that while the first root principle is not so flawed as

to be blithely dismissed, neither is it sufficiently strongly argued and defended to bear the weight of validating other principles. As we shall see, William immediately attempts such an invocation in defense of the second root principle.

2. The argument from incorruptibility

Another "root principle found among the philosophers," writes William, is that "every substance, whose form is not corruptible, is incorruptible."[16] What reasons have we to think that the soul is not corruptible?

First, "no intelligent substance has any of the material forms as its own, as natural and essential."[17] The reasoning here is the Aristotelian psychophysics whereby "a substance receives the likenesses of all forms" without the material.[18] "Since, then, only a material form is corruptible

---

[16]Ibid., 31.

[17]Ibid.

[18]Cf. Aristotle, *De Anima* III, 4,419a20-21:

> Therefore, since everything is a possible object of thought, mind in order, as Anaxagoras says, to dominate, that is, to know, must be pure of all admixture; for the co-presence of what is alien to its nature is a hindrance and a block: it follows that it too, like the sensitive part, can have no nature of its own, other than that of having a certain capacity. That that in the soul which is called mind (by mind I mean that whereby the soul thinks and judges) is, before it thinks, not actually any real thing.

and an intelligent substance cannot have that kind of form as natural or essential, it is clear that an intelligent substance is incorruptible."[19]

The objection is made that immaterial things do experience an immaterial corruption. Pains and torments, such as anger and envy, must cause injury, and if injury is possible so too is the extremity of injury, failure and death.[20] As in the objection to the first root principle William differentiates between two kinds of injury, one which is essential, the other not. Injury, he claims, is an equivocal term which cuts either way. In the case of mental injury, however, the damage is not to the essence of the soul, but to

> an external possession. And thus as the injury or destruction of the thing does not touch the essence of its owner, so the injury of this sort is not essential. And hence it does not bring about essential injury or failure.[21]

The emotional injury is not essential, he claims, because it inheres to a possession of the soul (i.e., the body) and not the soul itself. Were the soul dependent upon the body in any way, though, this would be a different story.

> Nothing in the animal soul prevents an emotion of this sort from becoming so violent that the death of the body results from it and, as a result the failure of the brute soul. *This would likewise happen with regard to the human soul, if its*

---

[19]William, *The Immortality of the Soul*, 31-32.

[20]Ibid., 33.

[21]Ibid.

> *essence depended upon the essence of its body* [as is the case with animals].[22]

In other words, the deflection of this objection depends upon the invocation of the earlier root principle, the Argument from Dependence.

By William's own account, the objection to the second root principle is fatal if the first root principle cannot successfully be invoked. However, we found that this first principle is tenable only if we arbitrarily ignore select passages in favor of others. As such, it lacks the firm logical foundation required to buttress another argument.

In summary, the viability of the second root principle of incorruptibility is built upon the foundation of the first principle. But because this latter is only weakly argued, it cannot support the weight of the argument for the former, and thus the argument's structure collapses.

3. The argument from natural desires

William's third root principle is that "no motion of nature is naturally pointless or in vain and that nothing that is naturally moved is kept by a natural impossibility from the end toward which it is moved."[23] This principle is surely meant to recall Aristotle's extensive considerations about motion.

---

[22]Ibid. (emphasis added).

[23]Ibid., 34.

For Aristotle each of the elements tends toward its natural resting place.

> The natural motions … are those motions proper to objects in virtue of their being the type of thing (elementally) they are — all the elements (earth, water, air, and fire) have internal, irreducible nisuses to move toward their natural places, which are determined absolutely: the heavy tend toward the centre of the cosmos…, while the light tend toward the extremities…. Thus the natural place for an element will be that region of space which is closest to the defining position which is capable of containing all of the element in question….[24]

Like William, Aristotle stipulates that the elements will tend toward their natural places unless they are interfered with.

> The elements will, other things being equal, occupy their natural places; and they will fail to move to them only if they are in some way prevented from doing so.[25]

Natural motions are neither pointless (the goal is to move toward its natural resting place) or in vain (if unimpeded, the motion will terminate in the successful arrival at the natural resting place).

Clearly William is invoking the Aristotelian physics. The difficulty is that despite the similarity in language, his actual application of the principle of natural motion is nothing at all like Aristotle intended.

---

[24]R.J. Hankinson, "Science," in *The Cambridge Companion to Aristotle*, edited by Jonathan Barnes (New York: Cambridge University Press, 1995), 141-142.

[25]Ibid.

William wishes (a) to sever the conclusion about natural motion from its premises about natural bodies; and (b) change the thesis of physical movement to natural place to one about purposive movement toward natural states. He thus invokes as validating authority an Aristotelian principle he rewrites, based on a demonstration he rejects.

(a). Can nonphysical things have motion?

William extends the category of natural motion to include categories which are not strictly physical, such as emotions, parrying the objections which challenge this extension by analogy rather unconvincingly. He observes that he means by "motion only the disposition which by itself is the way of attaining something else." Fear and flight fall into this expanded understanding. Motion is for William what it never was for Aristotle, a mental *intention* to attain some *state*.

Aristotle, however, rejected the possibility that psychic attributes can be motions:[26] "any being that undergoes change is a material being. We cannot prise these two things apart, even in thought, without incoherence."[27] The difficulty is that each of his identified species of

---

[26]Charlotte Witt, "Dialectic, Motion, and Perception: *De Anima* Book I," in *Essays on Aristotle's De Anima*, edited by Martha C. Nussbaum and Amélie Oksenberg Rorty (Oxford: Clarendon Press, 1992), 179-180.

[27]Martha C. Nussbaum and Hilary Putnam, "Changing Aristotle's Mind," in *Essays on Aristotle's De Anima*, edited by Martha C. Nussbaum and Amélie Oksenberg Rorty (Oxford: Clarendon Press, 1992), 32.

motion — locomotion, alteration, diminution, and growth[28] — "involves place (*topos*). [They] all occur somewhere, at some location. But, if one of these forms of motion is proper to the soul, then soul must have a location or place, and Aristotle thinks that this idea is implausible."[29] All motions are characteristic only of the elements, and are therefore limited to physical bodies. "It is always with respect to substance or to quantity or to quality or to place that what changes changes."[30] Souls possess none of these things, and hence "it is an impossibility that movement should be even an attribute of it."[31]

What Aristotle would deny, William will assert. The model of natural motion which Aristotle developed for and restricted to physical bodies, William extends to the mental and intentional. As the elements tend toward natural *place*, the human soul has an analogous natural

---

[28]Aristotle, "On the Soul," 406a12-13. A different list is given in the *Categories* (translated by E.M. Edghill, in *The Basic Works of Aristotle*, edited by Richard McKeon (New York: Random House, 1941), 15a13: "There are six sorts of movement: generation, destruction, increase, diminution, alteration, and change of place [locomotion]."

[29]Witt, "Dialectic, Motion, and Perception," 173-174.

[30]Aristotle, "Physics," translated by R.P. Hardie and R.K. Gaye, in *The Basic Works of Aristotle*, edited by Richard McKeon (New York: Random House, 1941), III, 1, 200b33-34.

[31]Aristotle, *De Anima*, I, 2, 406a2. If in this text William would refute this position, in his own *De Anima* he unambiguously endorses it. There "he concludes that... there can be no motion of an incorporeal soul." Moody, "William of Auvergne," 44.

state, which is "true and integral happiness."[32] Again invoking Aristotle's model, this happiness must be attainable unless unnaturally impeded. Furthermore, the nature of such a happiness must be permanent, otherwise it "will not be happiness or true immunity from misery. For whatever is subject to death is not happy; indeed, it is for that very reason wretched, because it is subject to extreme misery."[33] The soul, then, must be immortal.

(b). Can conclusions about "natural place" be analogously extended to the idea of "natural state"?

William replaces movement toward natural place with the thesis of natural fulfillment.

> And since spiritual things are acquired, just as bodily ones are, it is likewise necessary that there be, for those who acquire them, ways of attaining them in a fitting manner. These ways cease when they have been obtained, just as in the attainment of bodily things.[34]

By analogy with Aristotle, if flight is a spiritual motion, then the rule says that flight is *necessarily* successful unless *unnaturally* impeded. Somehow, though, this does not seem intuitively correct. If the shark swims faster than you do, then it seems obvious both that flight will fail and that there is nothing unnatural about the failure. The implication

---

[32]William, *The Immortality of the Soul*, 35.

[33]Ibid., 38.

[34]Moody, "William of Auvergne," 35.

would seem to be that we would have no desires unless we also had the means to fulfill them; wanting to escape the shark entails having the means to do so.

The nub of this problem lies in what it means for something to be "in vain." Despite its importance, William depends entirely upon Aristotle's reputation to support this claim. Yet he has so altered the basis for the claim that it can no longer be justified on that basis. What was a rational, logical assertion in the original Aristotelian setting is here ad hocery.

The editorial commentary on this passage claims that "William clearly rests his case upon the Aristotelian principle that God and nature do nothing in vain," citing *De Caelo* I, 4, 271a33.[35] But an inspection of the cited passage reveals a somewhat different assertion from Aristotle. The specific sentence is that "God and nature create nothing that has not its use",[36] a substantially weaker claim that what William needs.[37] A

---

[35]Ibid., 34, note 27.

[36]Aristotle, "De Caelo," translated by J.L. Stocks, in *The Basic Works of Aristotle*, edited by Richard McKeon (New York: Random House, 1941), 404.

[37]A consideration of the sentence in its fuller context only renders its appropriateness to William's needs all the more unlikely. Aristotle is here speaking about circular and rectilinear motions of physical bodies, which are altogether different from the kinds of emotional and psychological motions William includes.

body or thing can have its use without every motion being required to be expressive of a hidden destiny.

A similar claim is made in the *De Anima*:

> But animals must be endowed with sensation, since Nature does nothing in vain. For all things that exist by Nature are means to an end, or will be concomitants of means to an end. Every body capable of forward movement would, if unendowed with sensation, perish and fail to reach its end, which is the aim of Nature; for how could it obtain nutriment.[38]

One interpretation of this passage is that Aristotle meant "that nature does not put any creature in a position which calls for a certain faculty without also providing the faculty."[39] William's argument might thus be that "true and lasting happiness" could not be a natural desire for the human soul (and he takes it as a given that this is true) without the ability to attain this condition, requiring immortality. But immortality is not a faculty, and hence again a seemingly helpful passage from Aristotle aids William not at all.

So we can conclude that William is out on this particular limb on his own, despite any intimations that he is here relying upon established philosophical precedents. The problem here is not that he disagrees with Aristotle, but that he hopes to use an Aristotelian conclusion without the supporting philosophical justification; that is, he hopes to apply (b) in the

---

[38]Aristotle, *De Anima*, 434a30-434b2.

[39]Ross, *Aristotle: De Anima*, 54.

novel context of (a). He is free to reject or revise the latter, but he cannot then adopt the former without new arguments.

In summary, this third root principle, if acceptable, could yield a consistent argument for the immortality of the soul. However William offers no reason why we should accept the principle in the first place, especially given its immediate counterintuitive implications. He intends that we should accept the principle based upon its supposed relationship to the arguments by Aristotle, but the two theses are only superficially related.

We must conclude that William here fails to present the kind of rational arguments which he indicated would be his goal, and that the third root principle fails.

4. The argument from destructibility

William next identifies the only four means by which something can be destroyed; if none of these ways can destroy the soul, then the soul is immortal. The four ways to destroy are: (1) through division of its form from its matter; (2) through division of its integral parts; (3) when the essence of what sustains it is destroyed; and (4) by the removal of its cause.[40]

The soul "cannot be destroyed in the first way, because it is pure form and an immaterial substance that is not itself composed with the

---

[40]William, *The Immortality of the Soul*, 46.

sort of composition that arises from matter and form."[41] In responding to the objection that the soul is actually "composed from matter and form," William asserts that the soul's form is "incorruptible" because it has no "contrary by which it might be corrupted." Using the Aristotelian image that the red eye cannot see red[42], likewise since "the intellective power cannot understand anything unless there is present to it either the thing itself or its likeness, it is clear it cannot understand its contrary and, therefore, cannot have a contrary."[43] The objection requires that the intellective power have a contrary, and thus that there be something not intelligible to it. "This statement is *not an explanation, but* madness, and

---

[41]Ibid.

[42]The faculty of sight depends upon the potential of the eye becoming red, but it cannot be potentially what it is actually, and thus it cannot see red if it is actually red and not potentially red. Cf. Aristotle, *De Anima*, III, 4, 429a19-22:

> Mind must be related to what is thinkable, as sense is to what is sensible. Therefore, since everything is a possible object of thought, mind in order, as Anaxagoras says, to dominate, that is, to know, must be pure from all admixture; for the co-presence of what is alien to its nature is a hindrance and a block: it follows that it too, like the sensitive part, can have no nature of its own, other than that of having a certain capacity.... For this reason it cannot reasonably be regarded as blended with the body.

"William accepts, with little change, the Aristotelian analysis of the physiology of sensation." Moody, "William of Auvergne," 47.

[43]William, *The Immortality of the Soul*, 47-48.

hence we should not argue against it further; one cannot assert the existence of what is unintelligible."[44]

In the next paragraph, however, William does admit the existence of one unintelligible thing: the intellect. Just as one cannot see one's own eye directly ("just as sight is invisible"), so too the intellect cannot think itself, and is therefore unintelligible.[45] The combined force of these passages is that (a) only intelligible things exist; (b) the intellect is unintelligible; yet (c) the intellect exists. Either (a) or (c) must be surrendered, given (b).[46] If (c), then the whole point of the argument, to prove the immortality of the soul, is defeated; if (a), then William has

---

[44]Ibid. The editor's gloss on this passage is helpful: "To say that the intellective power has a contrary is to say that there is something that is unintelligible. But to know that it is unintelligible involves knowing something about it."

It is unclear what impact such a conclusion has upon the idea of God itself. A transcendent God, it would seem, would necessarily be unintelligible, and as such according to William's argument, nonexistent. That obviously would not be a claim he would be happy with.

[45]Ibid., 48. This assertion can be compared with that from Aristotle. For him, divine thought can only be about thinking: "Therefore it must be of itself that the divine thought thinks (since it is the most excellent of things), and its thinking is a thinking on thinking." Aristotle, "Metaphysics," translated by W.D. Ross, in *The Basic Works of Aristotle*, edited by Richard McKeon (New York: Random House, 1941), Λ, 9, 1074b33-34. .

[46]The likelihood of surrendering (b) is small since it has the imprimatur of Aristotle.

failed to rebut the objection that the intellective power is both form and matter, and therefore subject to the first kind of destruction.

The serious flaw with this section, however, is, as Teske observes, that "it is not clear where in this argument William explicitly rules out the destruction of the soul by the removal of what sustains it, i.e., presumably by the removal of the body. He mentions this point...in passing. Earlier, of course, he showed that the human soul did not depend upon the body [i.e., the first root principle]."[47] We have already noted the shortcomings of his attempts to refute a dependency relationship between the body and the soul; if he intends for that argument to address the current problem of destructibility, we must conclude that the weaknesses found there transfer to its application here. Worse, of course, would be the conclusion that William simply failed to argue against this kind of destruction, in which case the argument fails outright for being incomplete.

In summary, at least two of the four identified means of destruction are not adequately eliminated in the case of the soul. The first refutation applies only if the soul is not both form and matter. But William's response to this objection raised contradictory claims about existence and intelligibility, and therefore cannot be judged adequate. No direct argument against the third refutation is made. If William intended

---

[47]William, *The Immortality of the Soul*, 51, note 70.

the results of his proof of the first root principle to apply in this context, then the weaknesses identified there apply here, and again the argument against this kind of destructibility cannot be deemed internally adequate.

5. The argument from contrast with the senses

In the prior set of arguments William tried to establish the immortality of the soul via analogy from the senses to the intellect: As the first, so the second. In his final set of arguments he makes exactly the opposite claim: the immortality of the soul can be proved based upon the contrast of the intellect with the senses. In other words, they are so different that the former must be immortal. The problem here is not that this analogy fails because the two (the senses and the intellect) fail to be analogous in some respects (e.g., the analogy, if extended, does not "hold"), and that William is alternately emphasizing where the analogy applies, and then where it does not. Rather, the problem is that the analogy either holds or break downs in the *same* respect for the various arguments.

It would seem that both sets of arguments cannot be valid: Either the intellective power is like the senses, or it is not, but it cannot be the case that both descriptions are true and equally useful in proving the same point. To the extent we have reason to accept one, we have that much cause to reject the other.

Consider the following: "We will say, then, that a sense is not applied to things that are sensed *without its being likened to them*. Thus touch is not applied to what is hot without becoming hot, and sight to what is bright without being illumined."[48] We saw this claim in the previous section. William's point there was that because this is true about senses, it is necessarily true about intelligence (the soul): "the intellective power cannot understand anything unless there is present to it either the thing itself or its likeness, it is clear it cannot understand its contrary and, therefore, cannot have a contrary."[49] And again, because the intellective power has no contrary, it cannot be destroyed in this way.

But now William reverses himself. After reminding us about the physiology of the physical senses which had earlier provided a model for the operation of the intellect, he asserts now that "the intellect behaves in the *opposite* manner in this respect."[50]

> For when we sense something hot we necessarily become hot, and when we sense something bright we are necessarily illumined, but when we understand something hot we do not in any way become hot, nor do we in any way become colored when we understand a color.[51]

---

[48]Ibid., 52 (emphasis added).

[49]Ibid., 47-48.

[50]Ibid., 52 (emphasis added).

[51]Ibid.

It is not necessary to decide which characterization is more sound. All we require is the conclusion that they cannot both be true accounts. On the one hand, if the intellect does not assume the forms of what it understands, then it *can* have a contrary, because it was the joint implications of the assertions that "it would not be able to receive its contrary or a likeness of it, just as whiteness is not able to receive blackness or a likeness of blackness,"[52] and that only intelligible things exist, which generated the conclusion that the intellect had no contrary and thus could not be destroyed in that way. If this is true, then the intellect is indeed, *contra* William, "affected by the intelligible thing or things."[53]

If, on the other hand, it is not true, then the argument to deny the existence of a contrary for the intellect fails. The net result here is that the arguments from contrast with the senses are strengthened, but the argument against destructibility is nullified. Again, it is not necessary here to ascertain which (if either) argument is sustainable. Our needs are met with the realization that the fourth and fifth arguments cannot both be maintained because they depend on contradictory models of how the intellect works.

---

[52]Ibid., 47.

[53]Ibid., 52, note 76.

B. Criterion Two: Kind of Soul

William explicitly criticizes Plato's *Phaedo* because the soul discussed therein is too general: the arguments

> are common to all species of souls so that they are also applicable to the brute and to the vegetative soul. It is clear that the existence of these souls after the body and apart from the body is pointless and utterly useless.[54]

From this passage we see that William recognized the importance of the fact that the soul proved to be immortal be one that we would be happy to claim as ours. Surprisingly, having begun with this crucial insight, he errs in exactly the opposite direction. Whereas Plato's soul was too expansive for William, his own is too emaciated.

At no point is the soul for William anything more than our intellective powers. Granted, some language can be found that is more expansive. When he speaks of the "immortality of rational souls"[55] the apologist could agree that it is the immortality of rational souls which concerns William. But this need not require that only the rational aspect of the soul is immortal. Indeed, the manuscript editor concludes that "With Avicenna, William viewed man as the soul. William even goes so far as to maintain that the soul is not a part of man, but the man."[56] If

---

[54]Ibid., 33-34.

[55]Ibid., 23.

[56]Ibid., 28, note 16.

that is true, then everything that we consider essential to "man" should be preserved in his immortal soul.

But later language denies such a fullness to William's conception of the soul. The passages quoted above speak of the "intellective power" of the soul, which is explicitly distinguished from the "sensitive soul."[57] Only the latter is claimed to be immortal. The question then becomes whether a concept of the soul which is restricted to the rational powers comports with our idea of what it is to be human. At the very least it would seem that the soul envisioned by William will have no contact outside of itself since it lacks the sensitive powers.

Most people would be disappointed to learn that only their intellect (but none of their emotions, and only few of their memories) survives death.[58] Consequently, to the extent that William has succeeded in

---

[57]Ibid., 52.

[58]Writing in the *De Anima*, William concluded that

> it is possible, and perhaps necessary, that many memories should perish, especially of particular things and of sensible things which are hidden away in the store-house of memory; and I say this because those memories which are in the soul, that is, whose signs are impressed in the soul itself, are not necessarily taken away from the soul by the death of the body. Those memories, on the other hand, which are merely impressed on the aforesaid cellule of the brain, and deposited in it as if in a thesaurus, are indeed lost, not only through the death of the body, but by wounds inflicted on this cellule even during life.

Moody, "William of Auvergne," 55-56.

demonstrating the immortality of a soul, it is not the kind of soul most aspire to preserve.

C. Criterion Three: Psychological Continuity

While nothing in *The Immortality of the Soul* suggests that William believed humans and animals have kinship ties, he never obscures the fact that, psychologically, we have much in common. We are "animals-plus," but "angels-minus":

> The holy angels are entirely separate and entirely independent of bodies, and the vegetative and sensitive souls [e.g., animals] are entirely impressed upon bodies and dependent upon them.[59]

The nature of this claim is clarified by Moody's analysis of a passage in the *De Anima.*

> Animal sensitivity is, quite literally, sensitivity of the body to its environment — it is that function which the animal soul carries out in the animal body, whereby the latter reacts as an organized whole to its physical situation. The process, in animals, begins and ends with motion, and is neither free nor conscious. It is because this constitutes the limit of the powers that belong to animal souls, that animals are not immortal; they have no active powers beyond those whose function can be fulfilled only in a body, and therefore survival of their souls in separation from their bodies would be supervacuous.[60]

In other words, "the brute and vegetative soul is a material form in this way, that is, dependent upon its matter both as regards its being and as

---

[59]William, *The Immortality of the Soul*, 44.

[60]Moody, "William of Auvergne," 51.

regards its activity."[61] Animals are literally animated matter, nothing less, but nothing more. The enumeration of their materially enabled abilities exhausts what animals are capable of. There is no "remainder" to survive death since there is nothing "extra" during life.

William does not deny that humans have this aspect of materially-enabled capacities, and in this way he accepts the principle of psychological continuity. His primary point in the contrast is to argue that we have some powers beyond these. He capitalizes on this dual nature of the human soul, describing it is "an example of something in between.... by an essential proportion, then, some of its powers are immortal, some are mortal."[62]

There is nothing evolutionarily incompatible with the image of the human as being the aggregation of an animal-like materiality with a nonmaterially-based intellective power. While he asserts nothing we would recognize as an evolutionarily relationship between humans and animals, his philosophical edifice also seems not to require anything which would refute such a position, such as a refusal to recognize any similarities at all between humans and animals.

---

[61]William, *The Immortality of the Soul*, 32.

[62]Ibid., 44. This claim alone is sufficient to demonstrate that William's soul is not the "man himself." If the man is a combination of both immortal and mortal aspects, then the immortal dimension alone will be significantly less than the whole, and something far short of what we mean by "man."

In summary, we can say that William's image of the soul does not offend an evolutionary interpretation of human emergence.

D. Conclusions

Just as William's *The Immortality of the Soul* is a hodgepodge of arguments defending the immortality of the soul, our conclusions about those arguments range widely. The first root principle is weakly defended from an opposing argument, rendering incomplete the development of this argument. Prematurely launched, however, the argument is enlisted at least twice (root principles two and four) to buttress other arguments. I judge that application unjustified, given the actual state of the first principle.

To the extent we have reason to accept principle four, to that extent we must reject principle five, since they are based upon mutually exclusive assumptions about the operation of the intellect.

The only identifiable problem with root principle three is that we have no reason to believe that it is true, and much reason to conclude otherwise. Anyone disposed to accept his assertion that just as (Aristotelian) physical elements have their natural motions, so to do psychological motions such as desires entail their ultimate fulfillments, could probably arrive at a rational belief in the immortality of the soul.

The only inherent flaw in William's evolutionarily compatible arguments is the tension between four and five. If principles one and

three could be strengthened (thereby improving both two and four), a neo-Williamesque defense of the immortality of the soul that is internally persuasive is not unimaginable.

One reason why such a reconceptualizing of this philosophical line will not take place is because we are (or should be) uninterested in the results. These arguments depend upon a characterization of the human soul as being pure intellect only, devoid of any sensory experience or memories which flesh out what we mean when we speak of a person surviving his or her own death.

CHAPTER V

DAVID HUME'S
OF THE IMMORTALITY OF THE SOUL (1777)

A. Criterion One: Theory Consistency

David Hume, the arch-skeptic, identifies three kinds of arguments for the immortality of the soul, the metaphysical, the moral, and the physical. His task will be to demonstrate the inadequacy of each of these strategies. We shall be considering only the metaphysical and physical arguments.

The moral arguments, like the first section of William of Auvergne's *The Immortality of the Soul*, strive to reach conclusions about the soul from premises about God, particularly that he is just. Aristotle calls these kinds of arguments "contentious", that is, "it starts from opinions that seem to be generally accepted, but are not really such."[1] It may be true that *if* God were just, we could derive important conclusions about the soul. But we have no reason to think that this condition holds, and Hume (not surprisingly) gives us none here. Such arguments are

---

[1]Aristotle, "Topics," translated by W.A. Pickard-Cambridge, in *The Basic Works of Aristotle*, edited by Richard McKeon (New York: Random House, 1941), 100b24-26. The two other kinds of reasoning are the *demonstration* "when the premisses from which the reasoning starts are true and primary" (100a27) and the *dialectical* "if it reasons from opinions that are generally accepted" (100a30).

therefore more apologetic than philosophic, and therefore fall outside the purview of the present exercise.

1. Metaphysical arguments

Hume introduces the first section with the observation that metaphysicians suppose that the soul is immaterial, it being impossible for thought to belong to a physical substance.

He rejects this position for two reasons. First, drawing upon the conclusions of his own *Treatise of Human Nature* which he invokes only surreptitiously as an unspecified "just metaphysics," Hume believes that the idea of a "substance" is itself "wholly confused and imperfect."[2] If the very idea of a substance is incoherent, then so too must be any argument built upon the sundry *kinds* of substances.

The relevant section in the *Treatise* is Book I, Section V: Of the Immateriality of the Soul. There he speaks of philosophers as

> those curious reasoners concerning the material or immaterial substances, in which they suppose our perceptions to inhere. I know no better method, than to ask these philosophers in a few words, *What they mean by substance and inhesion?*[3]

This claim is subject to the common requirement that all ideas find their ultimate source in an impression. Hume defines impressions as "all our

---

[2]David Hume, "Of the Immortality of the Soul," in *Writings on Religion*, edited by Antony Flew (La Salle, IL: Open Court, 1992), 29.

[3]David Hume, *A Treatise of Human Nature*, edited by L.A. Selby-Bigge (Oxford: Clarendon Press, 1896), 232.

more lively perceptions, when we hear, or see, or feel, or love, or hate, or desire, or will."[4]

> As every idea is deriv'd from a precedent impression, had we any idea of the substance of our minds, we must also have an impression of it; which is very difficult, if not impossible, to be conceiv'd. For how can an impression represent a substance, otherwise than by resembling it? And how can an impression resemble a substance, since, according to this philosophy, it is not a substance, and has none of the peculiar qualities or characteristics of a substance?[5]

Hume challenges these philosophers, then, to identify the impression giving rise to the idea of a substance. "This question we have found impossible to be answer'd with regard to matter and body;"[6] Hume doubts the outcome will be any more positive when asked of the soul.

Is substance "something which may exist by itself?" Not likely, since this definition "agrees to everything," and fails "to distinguish substance from accident."[7]

> Thus neither by considering the first origin of ideas, nor by means of a definition are we able to arrive at any satisfactory notion of substance; which seems to me a sufficient reason for abandoning utterly that dispute concerning the materiality and immateriality of the soul, and makes me absolutely condemn even the question itself. We have no perfect idea of anything but of a perception. A substance is

---

[4]David Hume, *An Enquiry Concerning Human Understanding*, edited by Antony Flew (La Salle, IL: Open Court, 1988), 64.

[5]Hume, *Treatise*, 233.

[6]Ibid., 232.

[7]Ibid., 233.

> entirely different from a perception. We have, therefore, no idea of substance.[8]

That which we have no idea of, can hardly be predicated to something else, and to believe otherwise, as do many philosophers, is surely "wholly confused."

The same "just metaphysics" also argues the following. All our knowledge derives from our experience (i.e., the impressions that we have). No experience warrants a conclusion of the nature of "cause" and "effect".[9] Consequently, we cannot know in any justifiable way that "matter, by its structure or arrangement, may not be the cause of thought. Abstract reasonings cannot decide any question of fact or existence."[10]

So the first set of metaphysical arguments — that the soul is an immaterial substance because thought cannot inhere to a physical substance — is flawed on two counts. We have no idea what a substance is, so the task to develop a typology of substances is meaningless, and in any event we have no reason (i.e., we have no experiences which justify) the conclusion that matter could not support thinking.

---

[8]Ibid., 234.

[9]See relevant discussion on Hume's epistemology, especially as it pertains to the problem of induction, in Part I of this thesis.

[10]Hume, "Immortality," 30.

Hume continues: Even if there were a thought-inhering immaterial substance we call the "soul," we still have no reason to expect this soul to be immortal. Whatever we might know of such a soul substance would be by analogy with that other substance, matter. Nature, Hume argues, uses matter "as a kind of paste or clay" to be molded as needed. Importantly, matter is recycled; it "dissolves after a time each modification; and from its substance erects a new form."[11] So, by analogy, "As the same material substance may successfully compose the body of all animals, the same spiritual substance may compose their minds."[12]

Thus ends Hume's treatment of the metaphysical arguments for the immortality of the soul. Has he been successful in dismissing them? I think so. Hume's rebuttals depend upon his contention that all our knowledge must be directly linked somehow with experience. The justification of this position was the task of both the *Treatise* and the first *Enquiry*, and seems on the whole to be convincing. Once this point has been granted, Hume's conclusions follow directly that we have (a) no basis for the idea of "substance," and (b) certainly no basis for the idea of a soul substance other than from analogy with the idea of physical substance. Thus, for Hume, we either have no immaterial soul at all, or

---

[11]Ibid.

[12]Ibid.

one whose substance is reused as is matter, meaning that it cannot be immortal.

2. Physical arguments

The physical arguments are those "from the analogy of nature."[13]

> Where any two objects are so closely connected that all alterations which we have ever seen in the one are attended with proportionable alterations in the other, we ought to conclude, by all rules of analogy, that, when there are still greater alterations produced in the former, and it is totally dissolved, there follows a total dissolution of the latter.[14]

In other words, if the mind and the body parallel one another in the little things, we should expect them to continue in that relationship even for the big ones.

Sleep in the body, for example, is "attended with a temporary extinction" of the soul.[15] When the body is weak in infancy, and strong in adulthood, the mind is likewise. They both share a "sympathetic disorder" during illness, and experience a common decay in old age. Given how the soul tracks the body in all these conditions, we should expect that when the body experiences the most extreme of dissolutions in death, the soul here also does likewise.

---

[13]Ibid., 35.

[14]Ibid.

[15]Ibid.

We have other analogous reasons for conclude against the soul's immortality. "Judging by the usual analogy of nature, no form can continue, when transferred to a condition of life very different from the original one."[16] Fish, for example, die when removed from this original condition of the water to one of air. Likewise,

> What reason then to imagine that an immense alteration, such as is made on the soul by the dissolution of its body and all its organs of thought and sensation, can be effected without the dissolution of the whole?[17]

As everything is "in common between the body and the soul," the existence of the one must depend on the other. To remove the soul from the body, consequently, is to destroy both.

A third analogy builds upon the observation that "nothing in this world is perpetual."[18]

> How contrary to analogy, therefore, to imagine that one single form, seemingly the frailest of any, and from the slightest causes, subject to the greatest disorders, is immortal and indissoluble?[19]

As nothing of our direct experience is perpetual, how likely is that that the most fragile of things we can imagine will possess this virtue? According to Hume, not very.

---

[16]Ibid., 36.

[17]Ibid.

[18]Ibid.

[19]Ibid.

In all these arguments Hume's intention is not to argue that the soul is mortal; rather, again drawing upon his conclusions of the *Treatise*, his point is that to the extent we can have any idea about these things at all, it can only be through analogy with that which we do know, the material world and our mind. These are the things we actually have experience of. As the soul is asserted to contain the latter, whatever pertains to the mind/body pair must apply to the soul/body pair. But everything we know of the mind/body interaction bespeaks the mortality of both. So: either the talk of a soul is nonsensical (the metaphysical arguments) or to the extent the soul is a legitimate idea, it is through its parasitic relationship with the mind, and that relationship says that the soul is not immortal. We have no other basis on which to ground talk of the soul, if we should be talking about it at all.

A final argument in this section is not one from analogy, nor is it properly a physical one. Like William of Auvergne, Hume invokes the maxim that "nature does nothing in vain." While for William this principle proves the immortality of the soul, for Hume it plainly demonstrates the opposite:

> Were our horror of annihilation an original passion, not the effect of our general love of happiness, it would rather prove the mortality of the soul. For as nature does nothing in vain, she would never give us a horror against an impossible event. She may give us a horror against an unavoidable event, provided our endeavours, as in the present case, may often remove it to some distance. Death is

> in the end unavoidable; yet the human species could not be preserved, had not nature inspired us with an aversion towards it.[20]

Hume's argument is that if death anxiety is an innate problem for humans this fact supports the conclusion that the soul is mortal. If the soul is immortal, then death is an "impossible event," and nature would never instill in us a dread of an event that would never happen. On the other hand, nature would wish make us afraid of events which are unavoidable (as death is) in the long term if that would motivate us to avoid it in the short term.

The evaluation of this argument first depends upon the likelihood that death anxiety is an endogenous psychological state, and not simply an epiphenomenon of other things like religious instruction or "the effect of our general love of happiness." The problem of death anxiety is described in more detail in the following section on Heidegger. An in-depth consideration of the problem from both psychological and philosophical perspectives suggests that death anxiety is indeed a natural problem for human beings.[21]

Second, the argument also depends upon the premise that "nature does nothing in vain." This statement is extremely puzzling coming from

---

[20]Ibid., 38.

[21]Extensive discussion of the problem of death anxiety is offered in James M. Donovan, *Defining "Religion": Death and Anxiety in an Afro-Brazilian Cult*, Ph.D. dissertation, Tulane University (1994).

Hume. His sword of extreme empiricism cuts both ways: If he can criticize ideas like "substance" because they lack direct experiential grounding, he can be similarly asked what is the experiential support for this claim that "nature does nothing in vain." Nothing I can think of, certainly. This argument, when made in the voice of David Hume, should be put aside. But this failing does nothing, as far as I can tell, to further the cause of the immortality of the soul because he is not invoking it to rebut a contrary argument.

Aside from this aberration, Hume's arguments for the mortality of the soul are sound.

B. Criterion Two: Kind of Soul

Hume does not, in *Of the Immortality of the Soul*, present what is perhaps his strongest refutation of the concept of the postmortem self-preserving soul. The problem is not that there is not a soul to preserve the self, but rather that there is no self to be preserved by the soul or otherwise. For these arguments we must turn to his *Treatise* and the first *Enquiry*.

Again invoking his mantra that all knowledge derives from impressions, he observes (in Part IV, Section VI of the *Treatise*, entitled "Of Personal Identity"):

> [The] self or person is not any one impression, but that to which our several impressions and ideas are suppos'd to have a reference. If any impression gives rise to the idea of the self, that impression must continue invariably the same,

> throe' the whole course of our lives; since self is supposed to exist after that manner. But there is no impression constant and invariable.[22]

In other words, identity over time is a conclusion we draw, and not an impression we receive, and hence we have no rational basis for its assertion.

Whence, then, this idea of the self? Even when we introspect we fail to find it.

> For my part, when I enter most intimately into what I call *myself*, I always stumble on some particular perception or other, of heat or cold, light or shade, love or hatred, pain or pleasure. I never can catch *myself* at any time without a perception, and never can observe any thing but the perception.[23]

The idea of the self is constructed piecemeal from our perceptions, and without these perceptions we would have no grist for our self-mill. Consequently, "were all my perceptions removed by death, and could I neither think, nor feel, nor see, nor love, nor hate after the dissolution of my body, I should be entirely annihilated."[24]

What the self is, then, is "nothing but a bundle or collection of different perceptions, which succeed each other win an inconceivable

---

[22]Hume, *Treatise*, 251.

[23]Ibid., 252.

[24]Ibid.

rapidity, and are in perpetual flux and movement."[25] The sense of identity is not a property we observe in this bundle, but a quality we ascribe to it. Identity "is nothing really belonging to these different perceptions, and uniting them together; but is merely a quality, which we attribute to them, because of the union of their ideas in the imagination."[26]

Hume is unequivocal and unrelenting in his denial of the existence of a self: "The identity, which we ascribe to the mind of man, is only a fictitious one."[27] If true, then the arguments about whether there is a soul capable of functioning as a postmortem survival mechanism is bootless. Again, the essay directed specifically at the question of the soul does not raise the problem of the legitimacy of the concept of the self.

---

[25]Ibid.

[26]Ibid., 260. A technical description of this process might be the following:

> [Consciousness] is *spatially multiple, yet effectively single at any one time. It is an emergent property of non-specialised groups of neurons (brain cells) that are continuously variable with respect to an epicentre*, where an emergent property is taken to be a property of a collection of components that could not be attributable to any single member of those components.

Susan Greenfield, "Soul, Brain and Mind," in *From Soul to Self*, edited by M. James C. Crabbe (London: Routledge, 1999), 112.

[27]Hume, *Treatise*, 259.

But since he does invoke the *Treatise* at other points of the essay, it is not illegitimate for us to consider other of that tome's arguments.

In the words of one scholar,

> Hume's discussion of personal identity is the best there is; no one can feel the same about the problem after reading it as he did before; and, like so much that Hume says, it is incisive, penetrating, and most unsatisfying.[28]

It is very difficult for the reader not to accept Hume's arguments here. Despite the tight structure of his arguments, many of us still hope that he is wrong. The hesitation in following Hume down this road is admittedly emotional and not philosophical: If we are so completely wrong on a matter so deeply intimate as whether or not the "I" is not an entity to be encountered by myself and others, but instead a fabrication constructed in my imagination; if we are so wrong on this, how can we think we are any more correct about matters more distant from us, to which we presumably have even less access? Hume unintentionally

---

[28]Terence Penelhum, "Hume on Personal Identity," *Philosophical Review*, 64(1955), 571. His high esteem for Hume's argument may be a rhetorical flourish intended to increase the significance of his own arguments outlining its flaws. Penelhum essentially finds that Hume's problems with the concept of the self depend upon a misunderstanding about how we group things together and label them the same.

> Put generally, whether the result is logically absurd, or logically possible, or logically necessary, if the two phrases "the same continuing *x*" and "several different *y*'s" are used of the same thing depends entirely on what nouns we use to replace *x* and *y*. It does not depend on the words "same" and "different" in themselves. (p. 582)

removes the one certain thing in which even a skeptic like Descartes could place trust. Leaving what?, we might ask.

Still, our conclusion must clearly be that Hume denies the reality of the personal self, and as such he trips over the second criterion.[29]

C. Criterion Three: Evolutionary Compatibility

Twice in this short essay Hume refers to animals in ways which indicate that he assumes a principle of psychological continuity. The first occurs in his discussion of premise that since thinking cannot inhere in physical matter, a soul of immaterial substance must exist. He argues a point which apparently he believes most of his readers will consider a *reductio ad absurdum*: Since animals "undoubtedly feel, think, love, hate, will and even reason,[30] tho' in an more imperfect manner than man....[are] their souls also immaterial and immortal?"[31] He plainly expects a negative answer here, with the implication that if this line of reasoning does not justify immortal animal souls, how could it justify immortal human souls?

---

[29]This does not mean, incidentally, that Hume is wrong, but only that he refutes the common intuition sense of self we stipulated in Part I. Quite possibly that common intuition is wrong; but that is not the problem with which we are engaged here.

[30]Compare this list of animal mental abilities with that given in the first *Enquiry* for humans: We "hear, or see, or feel, or love, or hate, or desire, or will." Hume, *Human Understanding*, 64.

[31]Hume, "Immortality," 31.

The second reference to animals builds upon this base: Despite the mental similarities, animal souls are presumed to be mortal. Indeed, the anatomical similarities are not any stronger than the mental, yet no one doubts the validity of analogies on these grounds. "The *Metempsychosis* [the transmigration of souls] is therefore the only system of this kind, that philosophy can so much as hearken to."[32]

Hume obviously exhibits unqualified acceptance for the principle of psychological continuity. He goes so far as to assert this principle to highlight an inconsistency in the traditional Christian doctrine that animals have no immortal souls. Given the principle, he says, the ground for conclusions about animal souls is no worse than that for human souls, so rationally one must reach the same result about both. That the form of Christianity he assumes does not display this consistency, he implies, is one more reason to reject its assertions about the soul.

D. Conclusions

Hume offers a theoretically consistent, evolutionarily compatible theory *against* the existence of a self-preserving soul. The only *a priori*

---

[32]Ibid., 36. The full passage preceding this sentence is:

> The souls of animals are allowed to be mortal; and these bear so near a resemblance to the souls of men, that the analogy from one to the other forms a very strong argument. Their bodies are not more resembling; yet no one rejects the arguments drawn from comparative anatomy.

problem with his philosophy is the radical extent to which he contradicts the common intuition that our sense of a self refers to something real. Strictly speaking, Hume falters on this middle criterion.

That said, I feel compelled to offer that Hume's failure on this criterion seems qualitatively different from the failures we have encountered in other authors. The difference hinges perhaps upon the fact that while other authors accept the first premise of the common intuition (that we have a self), they deviate on the precise characterization of that self. Hume alone rejects both premises.[33] He has therefore so changed the terms of the debate that it seems simplistic to simply conclude that he has failed to meet this criterion. There is a sense in which his alternative is so drastically different that it is actually acceptable. It does not hurt that his arguments are, at least on the surface, logically seamless.[34]

---

[33]This is not to say that Hume alone among all philosophers has rejected the idea of the self; only from among those we have examined here. Contemporary philosophers who join Hume in this denial, and often in terms which at least echo Hume's "bundle" theory, include Daniel Dennett (*Consciousness Explained* [Boston: Little, Brown, 1991]) and Galen Strawson ("The Sense of the Self, in *From Soul to Self*, edited by M. James C. Crabbe [London: Routledge, 1999]).

[34]At least, it does not seem to be inconsistent with the cursory overview we have given here. We can still hold out hope that a deep analysis of Hume's texts will reveal either a logical problem with the account of personal identity itself, or with the links which connect this discussion to the rest of the *Treatise*.

By accepting the first premise, those soul-theories have accepted the condition of being evaluated by the second. That is, having agree that there is a self, they have made themselves comparable on the dimension of the kind of self being defended. These theories, when they fail to capture the common intuition of what it is to be a self, are unproblematically judged to have failed the criterion.

To do the same to Hume, though, seems difficult. Since he disagrees with the common intuition that there is even a self to be evaluated, to fault him here may not be fair. The common intuition may in fact have begged an important question, which revealed itself only in Hume's work.[35]

In the end, we must abide by the conditions established in Part I so that all soul-theories are evaluated on the same criteria. The standard of self has nowhere else been problematic. Hume fails to meet

---

[35]Question-begging has been a contentious point in other aspects of Hume's philosophy. For example, Richard Purtill accepts C.S. Lewis' evaluation that Hume has so defined natural law that it necessarily precludes the possibility of a miracle. The problem with this is that Hume does not present the impossibility of miracles as a matter of definition of natural law, but as a necessary outcome of the limits of public testimony and personal experience. Hume "first answers, 'Yes,' to the question whether Nature is absolutely uniform; and then uses this 'Yes' as a ground for answering, 'No,' to the question, 'Do miracles occur." C.S. Lewis, quoted by Richard L. Purtill, "Defining Miracles," in *In Defense of Miracles*, edited by R. Douglas Geivett and Gary R. Habermas (Downers Grove, IL: InterVarsity Press, 1997), 66.

Here, of course, it is we who have begged the question by presuming that the self exists, inquiring only what kind of self that may

it strictly (he does, obviously, fail to provide a theory which captures the intuitive sense of self), but the failure is such that in another context this would not be a fatal shortcoming.

---

be. The decision we must make is whether Hume should be penalized for our preconceptions.

CHAPTER 6

MARTIN HEIDEGGER'S *BEING AND TIME* (1927)

The conceptual work of the concept "soul" is assigned within Martin Heidegger's phenomenology to the category of "Dasein," although conceivably he would object to this interpretation. His overt claim is that his analysis of Dasein is neutral on the question of post-mortem existence; a critical scrutiny of the requirements of his philosophy itself, however, requires a positive denial of any such survival. Heidegger's implicit denial of the personal survival after death fails as a general refutation because it also denies a critical link between man and animals, and therefore lacks the required sensitivity to evolutionary concerns.

A. Criterion One: Theory Consistency

The confrontation with death provides a useful entry into the hermeneutic circle that is Heideggerian phenomenology.

Despite unreflective assumptions to the contrary, our anxieties about death are easily shown to be broader than concerns about mere corporeal continuance. First, one can whittle away large chunks of the body without incurring death. Parts can be replaced with other anatomical parts or even mechanical devices compromising neither life nor sense of self. Even the brain need not be intact to preserve life, its

viability depending less on the preservation of any specific organic isolate than upon some minimal structure: "there is much [a man] can lose without losing the self."[1]

Second, if the death problem were solely or even primarily reframeable as a problem of bodily continuance, then as we master techniques which extend that continuance, to that same degree our anxiety over death should decrease. Since we have yet to achieve immortality, we expect some amounts of death fear to remain; but also since our medical technologies have vastly extended life expectancies, we should expect that our death fears to be observably less than they were. But this does not seem to be the case for two reasons.

First, levels of death anxiety as they are currently quantified do not seem to vary according to the technological achievements of synchronically compared culture groups.[2] Second, we have no reason to believe that death anxiety was greater in less technologically advanced eras.[3] On the contrary, some historical reports seem to suggest that

---

[1]Søren Kierkegaard, *The Sickness unto Death*, translated by Howard V. Hong and Edna H. Hong (Princeton, NJ: Princeton University Press, 1980), 55.

[2]Richard Lonetto and Donald I. Templer, *Death Anxiety* (Washington: Hemisphere Publishing Corporation, 1986).

[3]One apparent exception to this claim for the universality of death anxiety is given by L.R. Farrell, *Outline History of Greek Religion* (Chicago, IL: Ares Publications, 1974), 39:

earlier societies which lacked our scientific advantages (but which, significantly, shared most of our metaphysical assumptions) were if anything less vulnerable to excessive death concerns.[4]

Against this background, where does the self reside to be extinguished by death? Not, apparently, with the body itself, even if it is in some sense centered on that body. These considerations force the conclusion that the living self is not a property of the body in the sense usually assumed, such that it is the death of the body which presents the philosophical problem. The problematic death *may* be coterminous with the body, so that when the body dies this other thing dies also. But even if co-occurring, the outcomes are distinguishable. Heidegger would

---

> there is no proof that the primitive mind of the Hellene brooded much on the problem of death, or was at all possessed with morbid feeling about it; and in pre-Homeric times he must have been freer from care in this matter than he was in the later centuries, if we accept the view of certain scholars that the elaborate ritual of 'Katharsis' or purification, which was mainly dependent on the idea of the impurity of death, ghosts and bloodshed, was wholly the creation of post-Homeric days.

What are we to make of such a claim? First, we note the precise assertion being made, not that the ancient Greeks were death anxiety-free, but that there is no "proof" that they were death anxiety-obsessed (i.e., "brooded" on it with a "morbid feeling"). The universality thesis depends on neither of these extreme positions.

[4]Philippe Ariès, *The Hour of Our Death*, translated by Helen Weaver (New York: Alfred A. Knopf, 1981).

not contradict these conclusions. The contrast will be between the possible interpretations of this non-corporealized death.

One obvious response is that framed by Kierkegaard. Kierkegaard defends the Christian belief in an immaterial spirit (or soul) which survives physical death, and that it is the death or life of this construct with which we are truly concerned.[5] The self now *contained* within the body is whisked away when that body expires, moving on to experience nonmaterial realities (i.e., God). Kierkegaard's readers (Ernest Becker,[6] for one) may not openly endorse this theological solution, but their focus on the body as a container for the self predisposes their work to allow for it or something very similar.

Yet the soul is only one way of accounting for the nonequivalence of the body and the living self, and not necessarily the best way. The immediate weakness is that rather than addressing the problem of death, it dissolves it. By thinking in terms of "spirit" the dilemma Kierkegaard and other spiritists claim to be resolving has become a straw man. How awful can death really be if there is an enduring spirit? Answer: not very. To assert the existence of an immortal soul in which the self survives intact is to deny the initial premise, that death is a uniquely

---

[5]George Hunsinger, *Kierkegaard, Heidegger, and the Concept of Death*, Stanford Honors Essay in Humanities, Number XII (Stanford, CA: Stanford University Press, 1969), 36-37.

[6]Ernest Becker, *The Denial of Death* (New York: Free Press, 1973).

"mortifying" psychological and philosophical conundrum.[7] In retrospect, these existentialists do not argue that death *is* a problem, but only that it *would be* a problem were it not for this enduring spirit.

Heidegger will also avoid an equation between the body and the self, but he will do so in a way which will both avoid supernatural consequences and preserve intact the horror of death.

As with other existentialists, the importance of death for Heidegger is that it forces a recognition of our finitude. In its face we are compelled to recognize our 'default' existential condition as *thrown*[8] and *fallen*.[9]

---

[7]The details of the problematic nature of death cannot be expounded here; the reader is referred to Becker's *The Denial of Death*. The point should be made, however, that death is itself the problem, not merely our knowledge about death, although certainly our lack of the one exacerbates the other. The source of our death anxiety is not merely our ignorance about what happens after death. Were that the case, philosophy could dispel the emotional state by addressing the cognitive one. While this insight would not lack impact, neither would it remove the problem altogether. Cf. James M. Donovan, *Defining "Religion": Death and Anxiety in an Afro-Brazilian Cult*, Ph.D. dissertation, Tulane University (1994). By analogy, understanding "hunger" does not take you very far in removing the pangs of starvation; so too does even a true philosophical comprehension of death in the abstract fail to dissolve the emotional trauma of the recognition of *personal* death.

[8]"Thrownness" refers to the fact that we are essentially part of an historical stream which we had no role in creating. Consider, for example, how we receive a language as a full-blown entity. Our job as children is to learn the language presented to us, not to create one de nihilo. In Heidegger's words,

> An entity of the character of Dasein is its 'there' in such a way that, whether explicitly or not, it finds itself [sich befindet] in its thrownness.

This confrontation calls us each to choose to become authentic, where "an authentic existence is one which makes manifest the possible *as* possible."[10] Only when we are aware that we have choices can we really make choices, and it is anxiety in the face of death that moves us from our complacent existence among the They (our fallenness).

Like Kierkegaard, Heidegger realized that "the concepts of man, time, and ultimate reality all intersect in the concept of death."[11] But the 'man' whose death concerns Heidegger is not 'man' in Kierkegaard's sense of biological entity. Instead of a metaphysical 'man' localized in the body, Heidegger grounded his analysis in the scrutiny of Dasein.

Martin Heidegger, *Being and Time*, translated by John Macquarrie and Edward Robinson (New York: Harper & Row, 1962), 174.

[9]"Fallen" refers to a "basic kind of Being which belongs to everydayness" (Ibid., 219), and indicates the state of "absorption in [...,] the character of Being-lost in the publicness of the "they" (p. 220). To be fallen is to exist in a state of inauthenticity, to have fled from the freedom of personal choices into the performance of predetermined social roles as they are provided from our thrownness.

[10]Charles Guignon, "Truth as Disclosure: Art, Language, History," *Southern Journal of Philosophy*, 28(supp., 1994), 105, 120.

[11]Hunsinger, *Kierkegaard, Heidegger, and the Concept of Death*, p. 72. While this essay underscores the differences between Kierkegaard's and Heidegger's response to the death problem, the many points of similarity between them are reviewed by Dan Magurshak, "The Concept of Anxiety: The Keystone to the Kierkegaard-Heidegger Relationship," in *International Kierkegaard Commentary, Volume 8: The Concept of Anxiety*, edited by Robert L. Perkins (Macon, GA: Mercer University Press, 1985), 167-195.

Whether interpreted materialistically (so that the self dies with the body) or spiritually (so that the self survives the body), the thesis that the self is co-terminous with the body is one Heidegger categorically rejects. Heidegger describes human existence as a "happening"[12] entailing interaction and relations. "My being — who I am — is nothing other than what unfolds in the course of my interactions with the world over the course of my life."[13] The necessity of these relations means that there is no self apart from the world: "Being-in is not a 'property' which Dasein sometimes has and sometimes does not have, and *without* which it could *be* just as well as it could with it. It is not the case that man 'is' and then has, by way of an extra, a relationship-of-Being toward the 'world' — a world with which he provides himself occasionally."[14] To be at all is always to-be-there, the German word for which is *Dasein*.

> Heidegger's reasons for introducing the term "Dasein" — which translated literally, simply means "there-being" — where it would seem natural to talk instead of about human beings, are manifold. First, in everyday German usage, this term does tend to refer to human beings, but primarily with respect to the type of Being that is distinctive of them; it therefore gives his investigation the right ontological ring. Second, it permits him to avoid using other terms that

---

[12]Charles B. Guignon, "Introduction," in *The Cambridge Companion to Heidegger*, edited by Charles B. Guignon (Cambridge, UK: Cambridge University Press, 1993), 8.

[13]Ibid., 9.

[14]Heidegger, *Being and Time*, 84.

> philosophers have tended to regard as synonymous with "human being", and have concentrated upon to the point at which they trail clouds of complex and potentially misleading theorizing. Time-hallowed terms such as "subjectivity", "consciousness", "spirit" or "soul" could only be prejudicial to Heidegger's enquiry.[15]

Dasein "is not [*Homo*] *sapiens*."[16] More, "it should not be thought of as an object or thing in any sense."[17]

White[18] highlights this understanding of Dasein in her claim that "the death of a person and the existential death of Dasein" are not at all the same kind of problem.

> Discussions of Heidegger's notion of death assume that Dasein dies when it ceases to be actual and that this happens when a person undergoes physical death. Or, if they recognize that Heidegger calls death a 'way to be' and that for him death is a matter of 'Being toward death', at best they consider death to be a matter of how a person cares about his physical death. Both assumptions are mistaken. Dasein's 'death,' in Heidegger's technical sense of this term, is very different from a person's 'death' in the ordinary sense of the word.[19]

We have not yet discussed what might be meant by the "death of Dasein; that follows below. But the relational essence of Dasein contrasts with

---

[15]Stephen Mulhall, *Heidegger and Being and Time* (London: Routledge, 1996), 14.

[16]Ibid., 15.

[17]Guignon, "Truth as Disclosure," 106.

[18]Carol J. White, "Dasein, Existence and Death," *Philosophy Today*, 28(1984), 52-65.

[19]Ibid., 56.

the physical existence of the body, and therefore "death" cannot apply to the two in the same way. White's point that death of body is not death of Dasein is therefore well taken, albeit with certain caveats. Dasein is positioned between the two possible readings of prior existentialisms. It is not *merely* epiphenomenon of material mechanisms, nor is it an immaterial, independently and presumptively eternally existing spirit. Dasein is a third possibility between Marx and Descartes.

Contrary to both alternative perspectives, physical death is not necessary for the death of Dasein. Even "when Dasein explicitly addresses itself as 'I here', this locative personal designation must be understood in terms of Dasein's existential spatiality,"[20] and *not* its physical spatiality. Because Dasein is not the body but Being-in-the-world, death is not only that which attacks the body but also the world constituting Dasein.[21] Death threatens our connection with the world and not merely the body except insofar as the latter is construed in

---

[20]Heidegger, *Being and Time*, 155.

[21]This perspective should respond at least in part to questions Derrida directed toward Heidegger: "'What is death for a *Dasein* that is never defined *essentially* as a living thing?' What can 'life' and 'death' mean in the discourse of the 'ethical' once it is decisively divided from the realm of the merely 'biological'?" David Clark, "On Being 'The Last Kantian in Nazi Germany'," in *Animal Acts*, edited by Jennifer Ham and Matthew Senior (New York: Routledge, 1997), 183.

terms of the former. It is "existential death ... [which] turns the prospect of biological death into utter despair."[22]

The physical body is but one node in the network of relations which constitutes the self as defined by its "Being-in-the-world." The body may not even be the most important of the self's nodes. For example, an unusually charismatic personality may insinuate himself into another's self-defining network and displace the self's body as the central tendency within the relational network.[23] Such a person would regard a personal death for the benefit of the charismatic leader as actually being a form of self-preservation.[24]

It is the very breadth of the ego integrity which is our Being-in-the-world that makes our being towards death fraught with both nihilating threats[25] and immortalizing creativity.[26] Even in our most vital

---

[22]Karl Jaspers, quoted by Merold Westphal, *God, Guilt, and Death: An Existential Phenomenology of Religion* (Bloomington, IN: Indiana University Press, 1984), 101.

[23]A convincing account of the ability of charismatic personalities to overwhelm their audiences is given by Charles Lindholm, *Charisma* (Cambridge, MA: Basil Blackwell, 1990).

[24]From this perspective, the charismatic follower may be understood as being "fallen" in the extreme.

[25]Maddi refers to the "small deaths" each of us endures whenever we experience loss of control or meaninglessness. Salvatore R. Maddi, "Developmental Value of Fear of Death," *Journal of Mind and Behavior*, 1(1980), 85-92.

moments, we are dying somewhere in what should be understood as the disarticulation of our Being-in-*that*-world. "Dasein is dying as long as it exists."[27]

If physical death is not necessary for the death of Dasein, neither is it sufficient. In a curious passage Heidegger observes that "a deceased Dasein can historically still be alive in an eminent sense, perhaps more authentically than in the time in which the Dasein itself properly was."[28] Dasein here seemingly takes on the properties of a Cartesian soul at least to the extent that it is available for ongoing development after the host's physical death.

Seeing as a persuasive case can be made that Dasein is not coterminous with *Homo sapiens*, the interpretive possibility arises that after *Homo*'s physical existence ends, Dasein yet lingers. Such a dualism is absurdly out of place in Heidegger's thinking, but the puzzle remains how Dasein can achieve authenticity *after* death.

This line of thinking is fueled by locutions within the Heideggerian corpus which relegate Dasein to a property of the human being, rather

---

[26]Such as seeking immortality through creative works. Michael A. Slote, "Existentialism and the Fear of Dying," in *Language, Metaphysics, and Death*, edited by John Donnelly (New York: Fordham University Press, 1978), 69-87.

[27]Heidegger, *Being and Time*, 295.

[28]Martin Heidegger, *History of the Concept of Time* (Bloomington: Indiana University Press, 1985), 271.

than being the being (but not just the body) itself. In favor of the latter reading are claims such as "We accordingly designate the entity which we also call man as the entity which *is itself its 'there.'* With this, we first come to the strict formulation of the meaning of the term '*Dasein*'"[29] and "Dasein itself, ultimately the entitative things which we call men."[30] At other points, however, Heidegger asserts a weaker relationship than this original equation. The phrase "the Da-sein of human beings," for example, occurs in both his essays "On the Essence of Truth" and "On the Question of Being." If Dasein is but a property or possession of human beings, then it becomes thinkable that Dasein can continue to exist in some sense after the death of the human being.

Heidegger does not deny these possibilities; rather, he claims that his philosophy is agnostic in these particulars.

> If 'death' is defined as the 'end' of Dasein — this does not imply any ontical decision whether 'after death' still another Being is possible, either higher or lower, or whether Dasein 'lives on' or even 'outlasts' itself and is 'immortal'. Nor is anything decided ontically about the 'other-worldly' and its possibility, any more than about the 'this-worldly'.[31]

But this denial is disingenuous, at least when compared to the demands of his philosophy.

---

[29]Ibid., 253.

[30]Ibid., 202.

Death can perform the anxiety-provoking function required by Heidegger only if it is truly a boundary condition, beyond which is non-being. If we knew, or thought we knew that death was not a boundary, but only a threshold to a new kind of being (à la Kierkegaard), death could become an occasion for fear but not of anxiety.

Of course Heidegger is correct in the technical sense that this belief by Dasein is not tantamount to a metaphysical assertion that there is in fact nothing about death. Perhaps all that is necessary for Heidegger is that Dasein be convinced that this is the case, whether it is actually the case or not. But see the problem that presents: in order to achieve authenticity, to recognize the truth of his existence, the person must incorporate the recognition of the nothingness that is death. But if this is a false recognition (e.g., death is not a nothingness at all, but a happy eternity in the Elysian Fields), the precondition of truth is the acceptance of a falsehood. In its greatest moment of unconcealment and dwelling in truth, Dasein is fundamentally deceived. Under those conditions, it may even be incorrect to say that future dwelling-in-truth depends on a falsehood; more likely, while one can *correct* a false existential understanding and replace it with a true one, but one cannot *build upon* a false understanding to create a true one. Heideggerian

---

[31]Heidegger, *Being and Time*, 292. A similar disclaimer can be found in his *History* (p. 314).

authenticity becomes impossible even in principle unless the understanding of death is a nothingness is itself true.

We must conclude that despite Heidegger's overt claims of agnosticism about what is beyond death, his existential analytic of Dasein requires that there be actually nothing beyond death. This proposition cannot be merely the belief state of Dasein.

For the present discussion the points to be highlighted are these:

1. The work of the concept "soul" is assumed by Heidegger's concept of "Dasein."

2. The two categories differ in that while the Kierkegaardian soul is distinct from but co-extensive with the physical body, Heidegger's Dasein is similarly distinct but lacks physical extension.

3. Given the way in which Dasein is defined as the relational network of being in the world, Dasein is more fragile than the soul, being as it is more sensitive to noncorporeal changes which constitute the "little deaths" which go on all the time. It is unclear how much "change" must be endured before one Dasein is deemed "dead," and another one arisen to take its place.

4. Heidegger claims that his philosophy is agnostic about the survival of Dasein after physical death. The main of this argument requires that such possibility be forcibly rejected: There is no personal survival after death.

B. Criterion Two: Kind of Soul

We need linger on this criterion but briefly. We must recall the pretheoretical nature of the concept of the self that we wish to see preserved in the death-surviving soul. As discussed in Part I, the hope is that the "I" which makes me *me*, should be mirrored in the qualities ascribed to the soul. For this purpose we need not presuppose just what it is that makes me *me*; our sole point is that each of us does have a sense of a *me* which we defend mightily, and that whatever the soul may be, if it is properly conceived it will be such that I recognize myself in it.

The entire thrust of the phenomenological enterprise is that it takes just such pretheoretical experiences as the primary data for philosophical analysis. Consequently, Heidegger's description of Dasein is perfectly compatible with our beliefs about the self: it incorporates our intuitions that it somehow importantly relates to the body, but also that it is not restricted to the body. Dasein is that which acts, which relates, and not merely that which thinks.

In no trivial sense, it was designed to achieve just this friendly result. Heidegger successfully meets the second criterion.

C. Criterion Three: Psychological Continuity

The existential analytic of Dasein as performed within *Being and Time* is but a means to an end for Heidegger. His larger project is to rethink the ontological difference, that is, the difference between Being

itself and beings. The analytic is important for Heidegger only because Dasein "is ontically distinguished by the fact that, in its very Being, that Being is an *issue* for it."[32] Understanding of Dasein is the tool by which a renewed unconcealment of Being itself can be achieved.

In addition to Dasein, Heidegger expresses his thinking about the ontological difference via the interconnected ideas of disclosure, unconcealment, truth, and clearing. Most of these concepts are brought together in the following passage discussing Van Gogh's painting of peasant shoes:

> What is at work in the work? Van Gogh's painting is the disclosure of what the equipment, the pair of peasant shoes, *is* in truth. This entity emerges into the unconcealedness of its being. The Greeks call the unconcealedness of beings *aletheia*. We say "truth" and think little enough in using this word. If there occurs in the work a disclosure of a particular being, disclosing what and how it is, then there is here an occurring, a happening of truth at work.[33]

Being itself is for Heidegger an act of disclosure, of unconcealment, and therefore of truth, which takes place in the ontological clearing. Dasein has a special role in this process: "Dasein, as constituted by disclosedness, is essentially in the truth. Disclosedness is a kind of

---

[32]Heidegger, *Being and Time*, 32.

[33]Martin Heidegger, "The Origin of the Work of Art," in *Poetry, Language, Thought*, translated by Albert Hofstadter (New York: Harper & Row, 1971), 36.

Being which is essential to Dasein. *'There is' truth only in so far as Dasein is and so long as Dasein is.*"[34]

The attribute which endows Dasein with this special significance is its relationship to Language. "Language is nothing but a distinctive possibility of the very being of Dasein."[35]

Heidegger's reverence for pure language is undisguised: "It is language that tells us about the nature of a thing, provided that we respect language's own nature.... Man acts as though *he* were the shaper and master of language, while in fact *language* remains the master of man."[36]

Language for Heidegger comes first, precedes Dasein. "It is language that first brings man about, brings him into existence."[37] "[Words] and language are not wrappings in which things are packed for the commerce of those who write and speak. It is in words and language that things first come into being and are."[38]

---

[34]Heidegger, *Being and Time*, 269.

[35]Heidegger, *History*, 261.

[36]Martin Heidegger, "Building Dwelling Thinking," in *Poetry, Language, Thought*, translated by Albert Hofstadter (New York: Harper & Row, 1971), 146.

[37]Martin Heidegger, "Language," in *Poetry, Language, Thought*, translated by Albert Hofstadter (New York: Harper & Row, 1971), 192.

[38]Martin Heidegger, *An Introduction to Metaphysics*, translated by Ralph Manheim (New Haven, CT: Yale University Press, 1959), 13. A

> The word, the name, restores the emerging essent from the immediate, overpowering surge to its being and maintains it in this openness, delimitation, and permanence. Naming does not come afterward, providing an already manifest essent with a designation and a hallmark known as a word; it is the other way around: originally an act of violence that discloses being, the word sinks from this height to become a mere sign, and this sign proceeds to thrust itself before the essent.[39]

"Sounds do not acquire meaning; rather, it is the other way around: meanings are expressed in sounds."[40]

Given this power of language, Dasein's access to it is determinative, and constructs qualitative differences between it and non-Dasein. "Where there is no language, as in the being of stone, plant, and animal, there is also no openness of what is."[41]

The essential gap between Dasein and non-Dasein (which can initially be loosely read as man and animal respectively) is therefore the linguistic aptitude of the first. For Heidegger, language is the key to Dasein's being what it is, and therefore the inability of animals to have a language comparable to human language renders them ontologically alien.

---

taste of this understanding of language's relation to the constitution of the world can be had in the novel by Suzette Haden Elgin, *Native Tongue* (New York: Daw Books, 1984).

[39]Heidegger, *Metaphysics*, 172.

[40]Heidegger, *History*, 210.

[41]Heidegger, "The Origin of the Work of Art," 73.

> It is held that man, in distinction from plant and animal, is the living being capable of speech. This statement does not mean only that, along with other faculties, man also possesses the faculty of speech. It means to say that only speech enables man to be the living being he is as man. It is as one who speaks that man is — man.[42]

Language is not for Heidegger an incremental ability, to be had in lesser or greater degree, despite suggestive studies to the contrary.[43]

Heidegger doubts that the human can helpfully be understood to be an animal at all:

> [It] finally remains to ask whether the essence of the human being primordially and most decisively lies in the dimension of *animalitas* at all. Are we really on the right track toward the essence of the human being as long as we set him off as one living creature among others in contrast to plants, beasts, and God?....
>
> Metaphysics closes itself to the simple essential fact that the human being essentially occurs in his essence only where he is claimed by being. Only from that claim "has" he found that wherein his essence dwells. Only from this dwelling does he "have" "language" as the home that preserves the ecstatic for his essence. Such standing in the clearing of being I call the ek-sistence of human beings. This way of being is proper only to the human being....
>
> The human body is something essentially other than an animal organism. Nor is the error of biologism overcome by

---

[42]Heidegger, "Language," 189.

[43]An excellent empirical overview for vervet monkeys is provided by Dorothy L. Cheney and Robert M. Seyfarth, *How Monkeys See The World* (Chicago: University of Chicago Press, 1990), while an unabashed advocate for the reality of animal communication is to be found in Donald R. Griffin, *Animal Minds* (Chicago: University of Chicago Press, 1992). A contrary opinion is reached by Stephen Budiansky, *If a Lion Could Talk* (New York: Free Press, 1998).

> adjoining a soul to the human body, a mind to the soul, and the existentiell to the mind.[44]

Such a bold statement precludes any obvious or comfortable allowance for an evolutionary theory which posits the biological relationship of all animal species. On the contrary, he registers his distaste for this hypothesis when he speaks with disdain of "our scarcely conceivable, abysmal bodily kinship with the beast."[45]

This "Heideggerian abyss which separates us from animals"[46] is encapsulated in a final theme in Heidegger's analysis of Being-toward-death. Only Dasein can *die*; animals (or those that only have life) perish.[47] As Clark notes, in Heidegger's philosophy only "the human properly dies, whereas the animal simply ceases to live."[48]

This deceptively clear conclusion is obscured somewhat by Heidegger's own intimation that despite the overt contrary claims we

---

[44]Martin Heidegger, "Letter on 'Humanism'," translated by Frank A. Capuzzi, in *Pathmarks*, edited by William McNeill (Cambridge: Cambridge University Press, 1998), 246-247.

[45]Ibid.. 248.

[46]Jennifer Ham and Matthew Senior, "Introduction," in *Animal Acts: Configuring the Human in Western History* (New York: Routledge, 1997), 4.

[47]Heidegger, *Being and Time*, 284. See also Martin Heidegger, "The Thing," in *Poetry, Language, Thought*, translated by Albert Hofstadter (New York: Harper & Row, 1971), 178.

[48]Clark, "On Being 'The Last Kantian in Nazi Germany'," 171.

have reviewed, *Dasein* is not restricted to "human beings." He does this by implication in his phraseology, which in a lesser writer we might dismiss as being a lapse without significance. But given that Heidegger is excruciatingly precise in his language use, we cannot afford to be so blithe.

We had earlier reviewed instances where his wording made it appear that Dasein was synonymous with human being, as when he writes about "the question of human essence in a regard that assures us an experience of a concealed ground of the human being (of Dasein)."[49] Heidegger also implies that Dasein is a property or attribute of human beings (e.g., "The Da-sein of human beings").[50] But this could be true while still restricting this property or attribute to humans. Importantly, he also employs the phrase "human Dasein."[51]

If Dasein *means* human, then the phrase is redundant, and very atypical of Heideggerian writing. The implication thus seems to be that there are other kinds of Dasein besides human Dasein, opening the possibility for the existence of a non-human, possibly animal Dasein.

---

[49]Martin Heidegger, "On the Essence of Truth," translated by John Sallis, in *Pathmarks*, edited by William McNeill (Cambridge: Cambridge University Press, 1999), 143.

[50]Ibid., 148, 151.

[51]Cf. Heidegger, "On the Essence of Ground," 108, 110, 112.

If animal Dasein exist, then the categorical divide Heidegger explicitly constructs would seem to have exceptions. Either some animals are possessed sufficiently of language to qualify as Dasein, or language is not really the distinctive feature of Dasein.

In any event, Heidegger's philosophy in the main must qualify as anti-evolutionary.

D. Conclusions

Heideggerian phenomenology offers a consistent argument against the post-mortem survival of a self-preserving entity, be it denominated "soul" or "Dasein." At worst a tension exists between the structure of the philosophy, and Heidegger's own "dicta" about that philosophy. Admittedly, he explicitly denies in sidebar commentaries many of the entailments of his worldview.[52] But the locomotive of his arguments has

---

[52]A skeptic might ask whether, if the explicit denials are that obviously contradictory, might they not be "explained away"? Why take them seriously?

A similar situation arises in the study of David Hume. Although his "actual arguments and the facts he adduces are regularly highly critical of religion," he will, again like Heidegger, largely in asides state personal convictions contradictory to his philosophical expositions. J.C.A. Gaskin, "Hume on Religion," in *The Cambridge Companion to Hume*, edited by David Fate Norton (New York: Cambridge University Press, 1993), 319. Gaskin's rightly points to the dangers in the eighteenth century to those who openly espouse atheistic views. If that is the case, Hume's contradictory asides can be philosophically ignored. But

> there remains an impression…that Hume really was unwilling to deny the existence of God and all lesser supernatural agents in the unequivocal sense now conveyed

an inertia which will not be overcome even by its author. Because these comments — always in the informal context of a chatty aside to his readers — are not on the same level as his serious philosophical elaboration, no genuine inconsistency exists within the structure of the philosophy itself.

It would seem, then, that beginning with *Being and Time* we see the unfolding of a consistent and compelling argument against survival of the self after death, and one which also fulfills our second criterion of being about the kind of self we would wish to have preserved.

On the third prong of our investigation, though, Heidegger argues for an impermeable barrier between humans and animals based upon the latter's possession of a developed symbolic language. It is not immediately obvious that this aspect of Heidegger's thinking is a

---

> by the notion of atheism. It is as if he was too consistent a skeptic to pronounce positively on any "remote and abstruse subjects," atheism included. (p. 321)

So even in the case of Hume, when we have good reason to imagine that contradictory asides are deliberate concessions to mere social convention and therefore without import, we should be cautious is assuming that to be the case.

Whatever social conventions Heidegger recognized against putting forth atheistic arguments were surely weaker than those Hume endured. That fact tends to increase presumption that we should take Heidegger's contradictory asides philosophically seriously.

In any event, we should remember the text-driven nature of this project. Our analysis has largely been restricted to what the author has actually said, and not what we think he might have meant. By this standard, the "chatty asides" must be presumed to impart information

necessary component of the larger philosophical enterprise which we have already judged to be a consistent argument for the lack of self-preserving soul. It may be possible to demonstrate that this thread is independent of the one outlined above, and not a required anti-evolutionary consequence. These thoughts are expressed in Heidegger's later pieces, and it is commonly suggested that a real divide exists between "early Heidegger" and "late Heidegger." Particularly, late Heidegger seems to move beyond the existential analytic featured in *Being and Time* into new directions. If that is the case, then the likelihood increases of the severability of the consistent anti-soul argument from the anti-evolutionary implications.

As it stands today, however, we must judge Heidegger's philosophy as incompatible with a natural world understood from an evolutionary perspective.

---

the author wished us to have and which we have no justification to ignore.

PART III

CONCLUSIONS

CHAPTER VII

SHOULD WE BELIEVE OR NOT?

The question of whether the self, as commonly imagined, survives death is one by which humans have been exercised for all of our written history, and probably before even that. This thesis examines five attempts to rationally justify such a belief.

Three criteria were identified which any such theory must successful meet before it could claim even *prima facie* reasonableness. These were: (1) that the theory be internally consistent, both with itself and with the wider corpus of that writer; (2) that the theory meaningfully capture the intuitive sense of the self that we are desirous of preserving; and (3) that the theory not make claims which are incompatible with other parts of our knowledge, specifically with evolutionary theory and its implications for psychological continuity.

The result of our critical scrutiny of five influential soul-theories is given in the summary table. We shall not rehearse here the arguments laid out in the earlier sections. Instead, we shall note some patterns which reveal themselves in the Table.

No soul-theory examined passed all three criteria. This means that, rationally, the ball is still in play. The existence of a self-preserving soul cannot be rationally defended, but neither can it be rationally denied. Perhaps some theory we have not yet scrutinized will pass our

tests; all that will mean, of course, is that it is worthy of really serious examination. But if in the end all theories fare no better, we may be compelled to decide that this question has no rational answer either way. We may have to concede that the question must be settled on other than rational grounds, and that it would not be irrational to do so.[1]

Summary Results of Five Critiques of Soul-Theories

	Consistency	Kind of Self	Evolution
Plato	Fail	Fail	Fail
Aristotle	Fail	Fail	Pass
William	Fail	Fail	Pass
Hume	Pass	Fail	Pass
Heidegger	Pass	Pass	Fail

That said, we can note that the scales are not evenly balanced between advocates and opponents of the self-preserving soul. No soul-theory advocating a self-preserving soul passed the criterion of internal consistency. On the other hand, two of three theories refuting such a soul (Aristotle, Hume, Heidegger) did meet this minimal standard. The implication seems to be that the soul is easier to deny than to defend.

---

[1]For a discussion on the possibilities of a rational irrationality, see Derek Parfit, *Reasons and Persons* (Oxford: Clarendon Press, 1984), 12-17.

Hume's attempt seems particularly attractive, and is the one I would select for subsequent improvement.

Hume's only flaw was to fail to meet the criterion that his theory reflect the common intuitions about the self. But then, so too did every other philosopher except Heidegger. If this failure rate continues as the samples of soul-theories increases, it may reflect more poorly on the common intuitions than on the theories. Any change in this criterion, though, would necessitate a drastic reformulation of the problem of death which motivates the project to find a death-surviving soul, and therefore should be attempted with only the extremist of cautions. But were the criterion altered, this change would only significantly impact the evaluation of Hume: he, and only he, would pass all three tests. For this reason, I personally find Hume's to be the best theory on the subject we have seen.

The novel analysis offered by this thesis has been the criterion of evolutionary compatibility. Only Plato and Heidegger fail the need to be compatible with evolutionary theory in the form of psychological continuity. Even before Darwin made evolutionary thinking standard fare for serious thinkers, philosophers seem to have been at least instinctively sensitive to the problem of the relationship between humans and animals, and to have found the position that they are not inviolably distinguished to be the most defensible. This ease with which

evolutionary tenets can be incorporated even to the oldest of our soul-theories may be a surprise for some, and even a pleasant one.

Had this criterion not been applied, Heidegger's phenomenological theory of Dasein would have been the only theory to pass both existing criteria, rendering it the best of the five, not Hume's. The introduction of this dimension of analysis, then, makes a difference.

BIBLIOGRAPHY

Alexander, Richard D. *The Biology of Moral Systems*. New York: Aldine de Gruyter, 1987.

Annas, Julia. *An Introduction to Plato's Republic*. Oxford: Clarendon Press, 1981.

Appenzeller, Tim. "Darwin, Fit as Ever." *U.S. News & World Report*, January 24, 2000, 49.

Ariès, Philippe. *The Hour of Our Death*, translated by Helen Weaver. New York: Alfred A. Knopf, 1981.

Aristotle. *The Basic Works of Aristotle*, edited by Richard McKeon. New York: Random House, 1941.

Axelrod, Robert. *The Evolution of Cooperation*. New York: Basic Books, 1984.

Barnes, Jonathan. "Life and Work," in *The Cambridge Companion to Aristotle*, edited by Jonathan Barnes. Cambridge: Cambridge University Press, 1995, 1-26.

Basinger, David, and Randall Basinger. *Philosophy and Miracle*. Lewiston: Edwin Mellen Press, 1986.

Becker, Ernest. *The Denial of Death*. New York: Free Press, 1973.

Bowler, Peter J. *Theories of Human Evolution*. Baltimore: Johns Hopkins University Press, 1986.

Brink, Chris. "Verisimilitude: Views and Reviews." *History an Philosophy of Logic* 10, 1989, 181-201.

Budiansky, Stephen. *If a Lion Could Talk*. New York: Free Press, 1998.

Burnyeat, Myles. *The Theaetetus of Plato*, translated by M.J. Levett. Indianapolis: Hackett, 1990.

————. "Is an Aristotelian Philosophy of Mind Still Credible? A Draft," in *Essays on Aristotle's De Anima*, edited by Martha Nussbaum and Amélie Oksenberg Rorty. New York: Oxford University Press, 1992, 15-26.

Cheney, Dorothy L., and Robert M. Seyfarth. *How Monkeys See the World*. Chicago: University of Chicago Press, 1990.

Clark, David. "On Being 'The Last Kantian in Nazi Germany'," in *Animal Acts: Configuring the Human in Western History*, edited by Jennifer Ham and Matthew Senior. New York: Routledge, 1997, 165-198.

Code, Alan, and Julius Moravesik. "Explaining Various Forms of Living," in *Essays on Aristotle's De Anima*, edited by Martha Nussbaum and Amélie Oksenberg Rorty. New York: Oxford University Press, 1992, 129-146..

Darling, David. *Soul Search: A Scientist Explores the Afterlife*. New York: Villard Books, 1995.

Darwin, Charles. *The Origin of Species by Means of Natural Selection*. New York: Avenel Books, 1979.

————. *The Descent of Man, and Selection in Relation to Sex*. Princeton: Princeton University Press, 1981.

————. *The Expression of the Emotions in Man and Animals*, Third Edition, edited by Paul Ekman. New York: Oxford University Press.

Dennett, Daniel. *Brainchildren*. Cambridge: MIT Press, 1998.

————. *Consciousness Explained*. Boston: Little, Brown, 1991.

Donovan, James M. *Defining "Religion": Death and Anxiety in an Afro-Brazilian Cult*. Ph.D. dissertation, Tulane University, 1994.

————. "Reinterpreting Telepathy as Unusual Experiences of Empathy and Charisma," *Perceptual and Motor Skills*, 87, 1998, 131-146.

Eldredge, Niles. *Time Frames: The Rethinking of Darwinian Evolution and the Theory of Punctuated Equilibria*. New York: Simon and Schuster, 1985.

"Evolution May Not Be Slow or Random." *Science News*, January 29, 2000, 75.

Farnell, L.R. *Outline History of Greek Religion*. Chicago: Ares Publishers, 1974.

Frede, Michael. "Plato's *Sophist* on False Statements," in *The Cambridge Companion to Plato*, edited by Richard Kraut. Cambridge: Cambridge University Press, 1992, 397-424.

————. "On Aristotle's Conception of the Soul," in *Essays on Aristotle's De Anima*, edited by Martha Nussbaum and Amélie Oksenberg Rorty. New York: Oxford University Press, 1992, 93-108.

Gallop, David. "Introduction," in *Phaedo*. Oxford: Oxford University Press, 1993.

Gaskin, J.C.A. "Hume on Religion," in *The Cambridge Companion to Hume*, edited by David Fate Norton. New York: Cambridge University Press, 1993, 313-344.

Gould, Stephen Jay. *The Panda's Thumb*. New York: W.W. Norton, 1980.

————. , "G.G. Simpson, Paleontology, and the Modern Synthesis," in *The Evolutionary Synthesis*, edited by Ernst Mayr and William B. Provine. Cambridge: Harvard University Press, 1980, 153-172.

Greenfield, Susan. "Soul, Brain and Mind," in *From Soul to Self*, edited by M. James C. Crabbe. London: Routledge, 1999, 108-125.

Griffin, Donald R. *Animal Minds*. Chicago: University of Chicago Press, 1992.

Gross, Paul R., and Norman Levitt. *Higher Superstition: The Academic Left and Its Quarrels with Science*. Baltimore: Johns Hopkins University Press, 1994.

Guignon, Charles. "Truth as Disclosure: Art, Language, History." *Southern Journal of Philosophy*, 28(supp.), 1994, 105-120.

————. "Introduction," in *The Cambridge Companion to Heidegger*, edited by Charles B. Guignon. Cambridge: Cambridge University Press, 1993, 1-41.

Hackforth, R. *Plato's Phaedo*. Cambridge: Cambridge University Press, 1955.

Ham, Jennifer, and Matthew Senior. *Animal Acts: Configuring the Human in Western History*. New York: Routledge, 1997.

Hanegraaff, Hank. *The FACE that Demonstrates the Farce of Evolution*. Nashville: Word Publishing, 1998.

Hankinson, R.J. "Science," in *The Cambridge Companion to Aristotle*, edited by Jonathan Barnes. New York: Cambridge University Press, 1995, 140-167.

Heidegger, Martin. *Being and Time*, translated by John Macquarrie and Edward Robinson. New York: Harper & row, 1962.

————. *History of the Concept of Time*, translated by Theodore Kisiel. Bloomington: Indiana University Press, 1985.

————. *Poetry, Language, Thought*, translated by Albert Hofstadter. New York: Harper & Row, 1971.

————. *An Introduction to Metaphysics*, translated by Ralph Manheim. New Haven: Yale University Press, 1959.

————. *Pathmarks*, edited by William McNeill. Cambridge: Cambridge University Press, 1998.

Hume, David. "Of the Immortality of the Soul," in *Writings on Religion*, edited by Antony Flew. La Salle: Open Court, 1992, 29-38.

————. *A Treatise of Human Nature*, edited by L.A. Selby-Bigge. Oxford: Clarendon Press, 1896.

————. *An Enquiry Concerning Human Understanding*, edited by Antony Flew. La Salle: Open Court, 1988.

Hunsinger, George. *Kierkegaard, Heidegger, and the Concept of Death*. Stanford: Stanford University Press, 1969.

Hyde, Dominic. "Sorites Paradox," in *The Stanford Encyclopedia of Philosophy*, Spring 2000.

Kierkegaard, Søren. *The Sickness unto Death*, translated by Howard V. Hong and Edna H. Hong. Princeton: Princeton University Press, 1980.

Kim, Jaegwon. *Philosophy of Mind*. Boulder: Westview Press, 1998.

Knowles, David. "William of Auvergne," in *The Encyclopedia of Philosophy*, edited by Paul Edwards. New York: Macmillan, 1967, 303.

Lakoff, George. *Women, Fire, and Dangerous Things: What Categories Reveal about the Mind*. Chicago: University of Chicago Press, 1987.

Lindholm, Charles. *Charisma*. Cambridge: Basil Blackwell, 1990.

Lonetto, Richard, and Donald I. Templer. *Death Anxiety*. Washington, D.C.: Hemisphere Publishing, 1986.

Luria, Salvador E., Stephen Jay Gould, and Sam Singer. *A View of Life*. Menlo Park: Benjamin/Cummings, 1981.

Maddi, Salvatore R. "Developmental Value of Fear of Death." *Journal of Mind and Behavior*, 1, 1980, 85-92.

Magurshak, Dan. "The Concept of Anxiety: The Keystone to the Kierkegaard-Heidegger Relationship," in *International Kierkegaard Commentary, Volume 8: The Concept of Anxiety*, edited by Robert L. Perkins. Macon: Mercer University Press, 1985, 167-195.

Majoria, S. "Speaker Lectures about Darwinian Legacy to the Scientific Community." *The Reveille [Louisiana State University]*, October 6, 1998, 3+.

Matthews, Gary. "Augustine and Descartes on the Souls of Animals," in *From Soul to Self*, edited by M. James C. Crabbe. London: Routledge, 1999, 89-107.

Mikalson, Jon D. *Athenian Popular Religion*. Chapel Hill: University of North Carolina Press, 1983.

Miller, David. "Popper's Qualitative Theory of Verisimilitude." *British Journal for the Philosophy of Science* 25, 1974, 166-177.

Mithen, Steven. *The Prehistory of the Mind: The Cognitive Origins of Art, Religion and Science*. London: Thames and Hudson, 1996.

Moody, Ernest A. "William of Auvergne and His Treatise *De Anima*," in *Studies in Medieval Philosophy, Science, and Logic*. Berkeley: University of California Press, 1975, 1-109.

Morgan, Michael L. "Plato and Greek Religion," in *The Cambridge Companion to Plato*, edited by Richard Kraut. Cambridge: Cambridge University Press, 1992, 227-247.

Mulhall, Stephen. *Heidegger and Being and Time*. London: Routledge, 1996.

Nagel, Thomas. *Mortal Questions*. Cambridge: Cambridge University Press, 1979.

Noble, William, and Iain Davidson. *Human Evolution, Language and Mind*. Cambridge: Cambridge University Press, 1996.

Nussbaum, Martha C., and Hilary Putnam. "Changing Aristotle's Mind," in *Essays on Aristotle's De Anima*, edited by Martha Nussbaum and Amélie Oksenberg Rorty. New York: Oxford University Press, 1992, 27-56.

Parfit, Derek. *Reasons and Persons*. Oxford: Clarendon Press, 1984.

Penelhum, Terence. "Hume on Personal Identity." *Philosophical Review*, 64, 1955, 571-589.

Penner, Terry. "Socrates and the Early Dialogues," in *The Cambridge Companion to Plato*, edited by Richard Kraut. Cambridge: Cambridge University Press, 1992, 121-169.

Perry, John. *A Dialogue on Personal Identity and Immortality*. Indianapolis: Hackett, 1978.

Piotrowski, William K. "The Kansas Compromise." *Religion in the News*, 2(3), 1999, 10-12.

Plato. *Republic*, translated by G.M.A. Grube. Indianapolis: Hackett, 1992.

————. *Meno*, translated by G.M.A. Grube. Indianapolis: Hackett, 1976.

Poirier, Frank E. *Understanding Human Evolution*, Third edition. Englewood Cliffs: Prentice Hall, 1993.

Popper, Karl. *Objective Knowledge*. London: Oxford at the Clarendon Press, 1972.

————. "A Note on Verisimilitude," *British Journal for the Philosophy of Science*, 27, 1976, 147-164.

Purtill, Richard L. "Defining Miracles," in *In Defense of Miracles*, edited by R. Douglas Geivett and Gary R. Habermas. Downers Grove: InterVarsity Press, 1997, 61-72.

Robinson, T.M. *Plato's Psychology*. Toronto: University of Toronto Press, 1970.

Rorty, Amélie Oksenberg. "*De Anima*: Its Agenda and Its Recent Interpreters," in *Essays on Aristotle's De Anima*, edited by Martha Nussbaum and Amélie Oksenberg Rorty. New York: Oxford University Press, 1992, 7-14.

Roscoe, P.B. "The Perils of Positivism in Cultural Anthropology." *American Anthropologist* 97, 1997, 492-504.

Ross, David. *Aristotle: De Anima*. Oxford: Oxford University Press, 1961.

Salisbury, Joyce E. "Human Beasts and Bestial Humans in the Middle Ages," in *Animal Acts: Configuring the Human in Western History*, edited by Jennifer Ham and Matthew Senior. New York: Routledge, 1997, 9-21.

Salmon, Wesley C. "Should We Attempt to Justify Induction?" *Philosophical Studies*, 8, 1957, 33-48.

Sheler, Jeffrey L. "Hell Hath No Fury." *U.S. News & World Report*, January 31, 2000, 45-50.

Simpson, George Gaylord. *The Meaning of Evolution*. New Haven: Yale University Press, 1950.

Slote, Michael A. "Existentialism and the Fear of Dying," in *Language, Metaphysics, and Death*, edited by John Donnelly. New York: Fordham University Press, 1978, 69-87.

Smith, Robin. "Logic," in *The Cambridge Companion to Aristotle*, edited by Jonathan Barnes. New York: Cambridge University Press, 1995, 27-65.

Sorabji, Richard. "Soul and Self in Ancient Philosophy," in *From Soul to Self*, edited by M. James C. Crabbe. London: Routledge, 1999, 8-32.

Stanley, Steven M. *The New Evolutionary Timetable*. New York: Basic Books, 1981.

Stone, Gregory B. "The Philosophical Beast: On Boccaccio's Tale of Cimone," in *Animal Acts: Configuring the Human in Western History*, edited by Jennifer Ham and Matthew Senior. New York: Routledge, 1997, 23-41.

Strawson, Galen. "The Sense of the Self," in *From Soul to Self*, edited by M. James C. Crabbe. London: Routledge, 1999, 126-152.

Sullivan, Harry Stack. "The Illusion of Personal Individuality." *Psychiatry*, 13, 1950, 317-332.

Tattersall, Ian. *The Fossil Trail*. New York: Oxford University Press, 1995.

Taylor, C.C.W. "Politics," in *The Cambridge Companion to Aristotle*, edited by Jonathan Barnes. New York: Cambridge University Press, 1995, 233-258.

Thomson, Judith Jarvis. *The Realm of Rights*. Cambridge: Harvard University Press, 1990.

Thornton, Stephen. "Karl Popper," in *The Stanford Encyclopedia of Philosophy*, Spring 2000.

Tichý, Pavel. "On Popper's Definitions of Verisimilitude." *British Journal for the Philosophy of Science* 25, 1974, 155-160.

Volkart, E.H. "Human Nature," in *A Dictionary of the Social Sciences*, edited by J. Gould and W.L. Kolb. New York: Free Press, 1964, 306-307.

Westphal, Merold. *God, Guilt, and Death: An Existential Phenomenology of Religion*. Bloomington: Indiana University Press, 1984.

White, Carol J. "Dasein, Existence and Death." *Philosophy Today*, 28, 1984, 52-65.

William of Auvergne. *The Immortality of the Soul*, translated by Roland J. Teske. Milwaukee: Marquette University Press, 1991.

Wilkes, K.V. "*Psuche* versus the Mind," in *Essays on Aristotle's De Anima*, edited by Martha Nussbaum and Amélie Oksenberg Rorty. New York: Oxford University Press, 1992, 109-128.

Wilson, Edward O. *On Human Nature*. New York: Bantam Books, 1978.

Witt, Charlotte. "Dialectic, Motion, and Perception: *De Anima* Book I," in *Essays on Aristotle's De Anima*, edited by Martha Nussbaum and Amélie Oksenberg Rorty. New York: Oxford University Press, 1992, 169-184.

Wrightsman, Lawrence S. *Assumptions about Human Nature*, 2nd edition. Newbury Park: Sage, 1992.

VITA

James M. Donovan was born and raised in Chattanooga, Tennessee. As an undergraduate he attended the University of Tennessee at Chattanooga, where he was a Brock Scholar and chaired the InterFraternity Council. He graduated in 1981 with a double major in Humanities (with honors) and Greek and Latin. In 1985 he enrolled at Tulane University in the Anthropology Department, and in 1986 at Louisiana State University in the School of Library and Information Science. He completed the latter program in 1989. With a dissertation based upon fieldwork in Rio de Janeiro among Candomblé cult members and titled "Defining *Religion*: Death and Anxiety in an Afro-Brazilian Cult," he received his anthropology doctorate in 1994. With his master's degree in philosophy now completed, he is scheduled to begin law studies at Loyola University in the fall of 2000.

Professionally Dr. Donovan has worked with the Tulane University Law School Library in various capacities since 1985, ending as Head of Access Services. Civic involvements include appointments to the New Orleans Mayor's Advisory Committee on Lesbian, Gay, Bisexual and Transgender Issues, which he chaired in 1998-1999, and to the New Orleans Regional AIDS Planning Council. Significant publications have appeared in *Law & Sexuality*, *Michigan Journal of Gender and Law*, *Journal of Contemporary Religion*, *Law Library Journal*, *Ethology & Sociobiology,* and *Method & Theory in the Study of Religion.*