The Logic of Immortality: Plato’s *Phaedo* and the Argument from Affinity

*Andrew M. Bailey*

**Abstract**

In this paper, I offer an analysis of the Argument from Affinity in Plato’s *Phaedo*. I will proceed by first outlining the basic project of *Phaedo*, showing the broader context of the Argument from Affinity. Second, I will offer two competing formulations of the Argument from Affinity, and suggest various textual clues in support of both. The more plausible of these formulations will turn out to be circular in nature. I will then show how this broad circularity is consistent with a general Platonic framework and is best seen as intentional. That the Argument from Affinity begs the question (is circular) may suggest, in fact, that it is a logical image of one central theme in the dialogue: the cyclical and circular nature of life and death.

It is tempting for contemporary students of philosophy to read classical authors with a jaundiced eye. Under this critical framework, apparent discrepancies or rational wrinkles are pounced upon as logical fallacies and instances of unsophisticated and lackluster argumentation. This judgment is easy, but far too hasty on several counts. As Alfred Freddoso notes:

> …without the sort of systematic study that we are not generally trained for either linguistically or philosophically, we contemporary … philosophers are not in a position even to understand, much less to criticize intelligently, most of the work of… classical metaphysicians. For in order to grasp what these authors are saying, we must immerse ourselves in their works and, at least initially, humbly submit ourselves to their tutelage; but this is a project that most of us have neither the expertise nor the time nor the inclination to undertake (Freddoso 1998, p.124).

This paper is an attempt to take up part of Freddoso’s challenge with regard to one Platonic dialogue, *Phaedo*. One easy reading of what I shall call the Argument from Affinity in *Phaedo* labels the line of reasoning a patent textbook fallacy. This reading is mistaken, or
so I hope to show. Plato deserves a more nuanced reading, and this paper is an attempt to
give it. More specifically, I shall here argue that the circular form of the Argument from
Affinity in Plato’s *Phaedo* is an image of the structure of the afterlife as understood by the
character of Socrates.

I will proceed by first outlining the basic project of *Phaedo*, showing the broader context
of the Argument from Affinity. Second, I will offer two competing formulations of the
Argument from Affinity, and suggest various textual clues in support of both. The more
plausible of these formulations will turn out to be circular in nature. Instead of merely
leaving the matter there, however, I will then show how this broad circularity is consistent
with a general Platonic framework and is best seen as intentional. That the Argument from
Affinity begs the question (is circular) suggests, in fact, that it is a logical image of one
central theme in the dialogue: the cyclical and circular nature of life and death.

*Phaedo* contains, “in a setting where literary brilliance and perception match the pathos of
the situation described, a lengthy exposition of what have been called ‘the twin pillars of
Platonism’…” (Robinson 1995, p.21) Here, the character of Socrates and, indirectly, the
author, embark on an expansive project, to argue for immortality.¹ Faced with his own
imminent death, Socrates attempts to persuade his friends that his soul will live on past the
execution of its body. In this way, he deflates the despair of his friends with hope in the
afterlife:
I want to make my argument before you, my judges, as to why I think that a man who has truly spent his life in philosophy is probably right to be of good cheer in the face of death and to be very hopeful that after death he will attain the greatest blessings yonder. (Plato 1997b, 63e-64a)

A variety of arguments are marshaled to ground this hope. The first of these is the Argument from Opposites (Plato 1997b, 70c-72d). Socrates here reasons that all things that come to be come to be from their opposites. Just as the larger comes from the smaller and the stronger from the weaker, so also, he concludes, does the living come from the dead. If the premises of this argument are correct, of course, it also follows that the dead come from the living. The resulting conclusion is not merely immortality—it is an endless cycle of reincarnation. This point will recur elsewhere in the dialogue.

Next, Socrates presents the Argument from Recollection to establish the pre-existence of the soul (Plato 1997b, 72e-77a). Given the mind’s knowledge of some things “themselves” (the example he uses is the Equal Itself), the soul must have, at one time, existed prior to its bodily incarnation. There is no other plausible explanation of its knowledge of the Forms. Socrates here seems to rely upon an argument made in an earlier dialogue, the *Meno*; the argument suggests that learning really just is recollection, the remembrance of knowledge had prior to birth:
... if the truth about reality is always in our soul, the soul would be immortal so that you should always confidently try to seek out and recollect what you do not know at present—that is, that you do not recollect? (Plato 1997a, 86b)

But this move is clearly not sufficient. That the soul pre-exists the body does not imply that it will live on after the body’s death. What I call the Argument from Affinity comes immediately after Socrates and Cebes recognize this deficiency. It functions, then, as the “second half of the proof” (Plato 1997b, 77c) in Socrates’ cumulative case for the immortality of the soul, to show that the soul will continue to live after bodily death.

Put prosaically, the Argument from Affinity goes something like the following. All of reality is divided into two realms. As Timaeus does, let us call these Being and Becoming (Plato 1997d, 28a-b). This is not to suggest that the terms are univocally related to Socrates’ conceptual framework in Phaedo. They will provide, however, a convenient means of understanding the metaphysics and logic of immortality. Being is marked by immutability, permanence, eternality, unity, divinity, intelligibility, and the like, while Becoming is fraught with constant change, mutability, plurality, and death. Being is the realm of Plato’s heaven—of the Forms, Becoming, the home of sensory experience and material objects. The human soul has important features in common with entities in the realm of Being; “… the soul is most like the divine, deathless, intelligible, uniform, indissoluble, always the same as itself…” (Plato 1997b, 80b) and hence is best seen as itself being in Being—a stone’s throw away from the conclusion that the soul is immortal.

Andrew M. Bailey
The most important claim here is that “…the soul more resembles the invisible, unchanging, and eternal forms than it does the visible, changing, and perishable objects that we perceive in this world” (Bostock 1999, p.422). The soul belongs, somehow, in one category and not the other.

I shall now examine the argument in greater detail, with an eye for its structure. My goal in this section shall be to understand the form of the argument and what inferences it may rely upon. The Argument from Affinity could be understood in at least two broad ways. While both will turn out to cohere with the text, the first, which I shall call ‘A,’ is logically invalid, while the other, ‘B,’ is not.

Formulation A of the Argument from Affinity could be seen as proceeding thusly:

A1. Everything in Being has feature F.
A2. Socrates’ soul (or souls in general) has feature F.
A3. Therefore, Socrates’ soul (or souls in general) is in Being.2

Note that while the argument may be generalized souls full stop, the text indicates most clearly that it is about one soul in particular. A1 first looks at a class of things (those that are in the realm of Being) and notes that they necessarily have some feature in common—invisibility, indivisibility, or a “divine nature.” A2 notes that Socrates soul, or all human
souls also have this feature. Does it follow, then, that Socrates’ soul is properly thought of as Being rather than Becoming? Certainly not. In fact, the inference of A3, regardless of its truth, is invalid. That is, it commits a compositional fallacy (or affirms the consequent). Just because Socrates’ soul has a feature which other objects in Being happen to also have does not imply that Socrates’ soul is itself in Being. To make the inference of A3 valid, premise A1 must be strengthened. One means toward this end is to posit:

A1’. An object x has F if and only if x is in Being.³

That is to say, having some feature F is both a necessary and sufficient condition of membership in Being. A1 could handily save argument A from charges of logical invalidity. If true, it, with A2 implies A3, which in turn gets Socrates his conclusion, that the soul is immortal. But is there any reason to believe this strengthened premise? Suppose F stands here for invisibility—is invisibility a necessary and sufficient condition for some object to be properly thought of as ‘being in Being?’ Not in the mind of this author—and it seems that decisive counter-examples will not be difficult to find. Consider a particular that is not physical—a mental event, for example. A mental event is not visible (or even physical) on the accounts of dualism most favorable to the Platonic project. It is not, however, anywhere close to the unchanging Forms found in Being. Mental events are mutable by definition. Their nature is to be transitory, to change, to come and to go. I shall
consider later the ramifications of supposing other features such as indivisibility to be necessary and sufficient conditions of membership in the realm of Being.

Argument A, as I have construed it is radically implausible. It is either logically invalid or subject to (what I take to be) decisive refutation. We may be tempted to leave the project here, but I suggest that charity demands we examine the argument anew, looking for a form that is perhaps less difficult to swallow. Let us call this reimagination of the argument ‘B.’

Under this reading of the Argument from Affinity, Socrates presents his listeners with an exhaustive disjunction—the terms he uses are ‘invisible’ and ‘visible.’ Following the convention established in argument A and Timaeus, let us call these Being and Becoming. Argument B, then, begins with the premise that everything belongs to one of these two general ontological categories. Socrates asks,

> Do you want to assume two kinds of existences, the visible and invisible? Let us assume this. And the invisible always remains the same, whereas the visible never does? Let us assume that too. (Plato 1997b, 79a)

We may understand this more formally as the following claim:

B1. Everything is in Being or is in Becoming.
Making a hidden premise in argument A explicit, let us add the notion that to be in Being is to be immortal:

B2. Everything that is in Being is immortal.

Where Argument A and B differ, however, is in the next premise. It seems to me to best understand Socrates’ point as a denial of one disjunct in B1 with regards to a particular thing or class of things. If Socrates can give evidence that the soul does not belong in Becoming, he has prima facie evidence that the soul in fact belongs in Being. An example may be helpful—if I can demonstrate that every crow must be either black or white, and I furthermore demonstrate that some crow (call him Jim) is not black, I have constructed a valid deduction to the conclusion that Jim is in fact white. Following this form, Socrates’ premise is that the soul somehow does not belong in Becoming:

B3. Socrates’ soul (or souls in general) is not in Becoming.

What validly follows from these three premises is the following:

B4. Therefore, Socrates’ soul (or souls in general) is in Being
B5. Therefore, Socrates’ soul (or souls in general) is immortal.\(^4\)
This formulation of the argument, unlike our first pass at argument A, is valid—its conclusion follows deductively from its premises. This argument, furthermore, seems faithful to Socrates’ exchange with Cebes excepting for one point: Socrates and Cebes always discuss their premises of their argument in terms of probability and likelihood (Plato 1997b, 78b). The current formulation of argument B does not. This will not pose any threat to my current purposes for two reasons. First, argument B could easily be amended to accommodate probability—by simply appending “it is probable that…” before the claim made by each premise or conclusion. The internal structure or form of the argument and the inferences it relies upon remain the same. Second, to be fair to Plato and his character of Socrates, I will neither offer nor consider any criticism of argument B that depends upon the argument being categorical rather than probabilistic.

Since the conclusions of argument B are entailed by the premises, it is valid. If one has good reason to believe B1, B2, and B3, one has good reason to affirm B4 and B5—and Socrates will have succeeded in his task of arguing for immortality. The real question, then, is the truth of the premises; the success of argument B must revolve around an evaluation of each premise’s strength or reasonability.

B1 merely claims an exhaustive disjunction, one side of which every object must fall.

Upon reflection, it seems that B1 is consistent with the metaphysical framework discussed in *Phaedo* and in other Platonic dialogues (Plato 1997b, 102a-103b and Plato 1997c, 475-
476d). This counts as evidence in its favor. Socrates would seem to be “in his rights” to rely on a premise such as this, given that he has elsewhere argued extensively in its favor, or in favor of its supporting philosophical framework.\(^5\)

B2 is equally innocuous. At least three grounds suggest its truth. First, immortality could simply flow from the definition of ‘Being.’ To be in Being just is to have features such as immutability, impassibility, purity, and the like—immortality certainly “fits in with the crowd,” so to speak. Second, B2 follows as a consequence from a Parminidean argument which apparently had some influence on Plato’s thought, that what is cannot not be, so a thing having Being cannot cease to be—that is, die (Kirk, Raven, and Schofield 1983, p.239ff). Finally, inductive evidence could be rallied in favor of B2 from various examples of entities which are in Being. That every (or at least most) examples of things in Being have immutability is prima facie evidence in favor of B2. This route is precisely the one Socrates follows, when he asks, “… can the Equal itself, the Beautiful itself, each thing in itself, the real, ever be affected by any change whatever?”

When a valid argument has as controversial conclusions as B4 and B5 and opening premises as innocuous as B1 and B2, it seems obviously the case that the other premise is doing the lion’s share of the logical work. This suspicion will be confirmed in light of the following claim I shall argue for: B3 begs the question at hand. Before explicating why
this is the case, I shall consider the evidence in favor of B3 and what it is to beg the question.

Socrates supplies his audience with at least three lines of reasoning in support of B3: three reasons why the soul is not at home in the sensible world of Becoming. First, he observes that the soul is not divisible, while material things are composite (Plato 1997b, 78c). Second, he claims that the soul is confused by the objects of this world, which is supposed to be clear evidence that it does not belong here (Plato 1997b, 79c-e). Third, the nature of the soul is not to be ruled, but rather to rule (Plato 1997b, 80a).

As it shall turn out, each of these evidences only begs the question against Socrates’ dialectical opponent, Euchecrates. To clarify what I mean by this, consider the following definition of what it is to beg the question:

BQ: If one or more of argument Q’s premises’ plausibility depend upon the plausibility of Q’s conclusion, then Q begs the question.

This captures, I think, the basic intuitions on the matter. To put BQ in other words, if one must already find a conclusion plausible to find the premises plausible, the argument has begged the question. To charge an argument, then, with begging the question is not to charge it with invalidity or false premises. Instead, the accusation is better seen as
epistemic: an argument for the immortality of the soul that begs the question can give no independent justification for belief in the immortality of the soul. Consider Plantinga’s notorious example of the principle at hand (Plantinga 1974, 217):

\[
P1. \text{Either God exists or } 7+5=14, \\
P2. \text{It is not the case that } 7+5=14, \\
P3. \text{Therefore, God exists}
\]

If one believed P3, one would likely find P1 plausible. Evidence in support of the conclusion would imply the premise. And yet, if one did not already believe P3, P1’s disjunction would seem odd at best, and carry with it very little epistemic weight.

With this background work done, let us now consider the strength of B3 and Socrates’ various proofs in its favor. Could one have good evidence for believing B3 without prior commitment to B4 or B5? We must remember the dialectical plot and literary context of this argument in *Phaedo*; Socrates is attempting to persuade a group of agnostic (undecided) listeners that his brand of substance dualism is to be preferred over the emergentism of Euchocrates (Plato 1997b, 88b-c). Given the cultural and philosophical context, Socrates’ audience would find the immortality of the human soul more plausible than would, say, an audience of contemporary naturalists.\(^7\) We must not stretch this point too far, though, for his audience remains uncommitted to the doctrine.

Andrew M. Bailey
Would such an uninterested audience find plausible the first claim, that the soul is indivisible? It’s difficult to see how this could be the case. After all, if the listener is not already committed to the immortality of the soul, its being divisible, material, or corruptible in composition seems a live option. Absent any further argument from Socrates, he has begged the question to assume this point. The best grounds for believing the soul to be indivisible is that it is immortal—but this is precisely the point in question.

The same analysis applies to the second line of evidence for B3. One not already persuaded by the Platonic framework of soul immortality and succession is likely to be confused by the notion of abstract “things themselves” that Socrates is fond of referring to. To an audience such as this (analogous to a contemporary empiricist), talk of concrete particulars would be far more tangible and cogent than talk of universals and Forms.

So also, the claim that the nature of the soul is to not be ruled, but rather to rule, is equally suspect. If one were not inclined to believe that the soul is an incorporeal, immortal substance incapable of death, why would one believe that the soul is divine? In fact, the dialectical alternative before the audience of Phaedo, Euchocrates’ proposal, suggests that exactly the opposite is the case, and that the body, in fact, is the ruler, while the soul only emerges as a sort of supervening harmony.

Andrew M. Bailey
Whether Socrates’ audience is agnostic or previously committed to Euchocrates’ emergentism, none of his three proofs are likely to be persuasive. If one is persuaded by Socrates’ conclusion, the premises, of course, will seem plausible. But as we’ve seen, these premises will likely seem dubious to any one not already committed to their conclusion. The Argument from Affinity rests in tension, then, between a logically invalid form (argument A) and a question-begging form (argument B). Given space limitations, I cannot defend this claim in depth, but I suggest that any reformulation of the Argument from Affinity is subject to this tension.

If one were to map out the Argument from Affinity as I have conceived it, the premises would depend upon the conclusion, and the conclusion upon the premises, suggesting a broadly circular form. To label the argument as circular is not to claim it is a useless or bad argument, though. One need not be left with a dialectical stalemate just because an argument turns out to be circular. This conclusion, in fact, will turn out to have several substantive implications for interpreting the dialogue as a whole. I suggest that there are at least three.

First, the Argument from Affinity is a logical image representing the psychic cycle. Just as a painting can resemble a scene, so also, an argument’s form can resemble some phenomena. What the resemblance relation amounts to, or even what it amounts to merely in the Platonic corpus is no small question—and it is a question beyond the scope of this
paper—but the relation is certainly instanced in some artifacts, and among these are some arguments.

This is precisely what I suggest may be happening with the Argument from Affinity, when taken in its literary context. Phaedo is a work, not just of philosophy, but of literature—it is, after all, a dialogue. As such, it is replete with images.

The Myth of Tartarus is a literary image of the afterlife, presenting a vivid picture of what is to happen to souls after their separation from the body (Plato 1997b, 110c-112b). Every soul is subject to a cycle of life, death, and reincarnation. After arguing extensively towards this end, the Myth of Tartarus is presented, giving Socrates’ listeners a mental depiction of what has been discussed. If Socrates hopes to instill hope in his audience, this certainly seems a wise strategy. The Platonic soul, after all, has both rational and non-rational parts, and Socrates reaches out to these diverse parts and faculties by using both rational (i.e., an argument) and non-rational (i.e., a myth or story) means.

The Argument from Opposites, too, may be an image of the cyclical psychology Socrates hopes to persuade his listeners of. In its content, it is repetitive, with recurring premises. Its substance mirrors the eternal oscillation of the soul from one body to the next.

Given that Phaedo is replete with images, and that repetition and cycle are central themes of the text, it is reasonable to look for pieces and parts of the whole which reflect these
themes in form and not merely content. This, I suggest, is one function of the Argument from Affinity in the dialogue, as a circular and cyclical argument.

Second, I observe that circular arguments dovetail with Plato’s pet metaphor (and accompanying experiential theory) of knowledge—sight. A detailed defense of this claim is beyond the scope of this paper, of course, but I outline the point in broad strokes. In key passages through his corpus (most notably the image of The Cave in Republic VII), Plato chooses to speak of knowledge as seeing. Knowledge is grounded in an experience, much like sense perception is grounded in the process of seeing. To know x just is to see or to “feel” that x is the case. Knowledge is what happens when a cognitive faculty makes contact with some aspect of reality. True knowledge, furthermore, is incorrigible because of its causal origin and grounding in experience. I cannot but believe that the sun is bright when I look at it; there is simply no other option. This suggests the beginnings of an experiential analysis of knowledge such that a correct statement of the necessary and sufficient conditions for ‘S knows that p’ will involve some experience had by S.

Plato’s choice of dialogue as a writing medium is uniquely experiential. Unlike a lecture or a plainer philosophical essay, a dialogue draws its reader in through dramatic elements, characters, plot (though often minimal), and movement. Sayre makes this point quite well, so I shall quote him at length:

Andrew M. Bailey
There are various ways of describing this remarkable capacity of Plato’s dialogues to reward one’s reading… On the most superficial level, the dialogue will contain a series of less than clever opinions… On a more substantial level, the typical dialogue will begin the engage the reader’s own serious thoughts about the matter under discussion, perhaps leading him or her to think up alternative responses to the mater dialectician’s questioning. Thus an experienced reader of the first part of the *Parmenides*, for example, might be stimulated to ‘come to Plato’s assistance’ with more sophisticated responses to Parmenides’ arguments against the Forms… Then there is the level on which a person stands to get an inkling of what Plato is getting at in the dialogue and begins to see ramifications not articulated by individual characters… (Sayre 1995, 28)

Dialogues uniquely involve the reader by offering them an experience. In this way, they can supply us, not merely *arguments* for some conclusion, but can also help us to *just see* that the conclusion is the case by *showing* it. The use of myth and imagery in *Phaedo* is just such a device—it is one additional tool of persuasion in the hand of a master philosopher. The Argument from Affinity by *being* circular in discussing the afterlife *shows* the reader the very thing Socrates is after. The circular argument functions as an invitation to enter the cycle, by offering the reader or listener experiential grounds of belief. This theme is consistent with a thread of classical thought, running back to Pythagorus, in which circles symbolize just the concept at hand.

Third, the image of the Argument from Affinity dovetails further with Socrates’ notion that to do philosophy is to practice dying (Plato 1997b, 64a, 68a). Under his view, dying just is part of an endless cycle of motion. Much of the work done in *Phaedo* reflects this, by tying in eternal oscillation as a literary theme and philosophical premise. The Argument

Andrew M. Bailey
from Affinity, too, does this in its form. To do philosophy as we have done (i.e., to attempt to uncover an arguments form or structure) is to gain a window into the afterlife and its circularity. By undergoing an experience that structurally resembles that of the afterlife, we are imprinted with a mental image of that afterlife. Interpreting the dialogue in this way adds cohesion to its parts. With the notion that the afterlife is circular in structure and that (at least one of) the arguments pursued in the dialogue are circular in their structure, we can make further sense of Socrates’ cryptic remarks about the relationship between philosophy and death. This is a further point in favor of the interpretation I have argued for.

Whether Socrates’ (and, indirectly, Plato’s) project is a success is a matter beyond the scope of this paper. Nonetheless, the textual and philosophical analysis here justifies enhanced respect for Plato as an author deserving of detailed attention. That the dialogue maintains cohesion and continuity even under scrutiny is a virtue; and this is a virtue we cannot observe without subjecting it to the very scrutiny and analysis I have demonstrated. Adopting the interpretive framework I have argued for can only increase our understanding and appreciation of the dialogue. 9
REFERENCES


NOTES

1 I shall assume in this paper that the dialogue is about what it appears to be about, viz., immortality. Readers interested in an alternative interpretation may consult Bostock 1986.

2 With B standing for the predicate ‘being in Being,’ we may render argument A formally as:
   A1. (∀x)(Bx ⊃Fx) For every x such that x has Being, x has feature F
   A2. (∃x)(Fx) There exists some x (Socrates) such that x has feature F
   A3. ∴ (∃x)(Bx) Therefore, there exists some x (Socrates) such that x has Being (A1, A2)

3 Formally rendered, that is:
   A1’. (∀x)(Fx ↔ Bx) For every x, x has feature F iff x has Being

4 Argument B as I have envisioned it can be rendered formally as follows:
   B1. (∀x)(Bx v Cx) For every x, either (exclusive) x has Being or x has Becoming
   B2. (∀x)(Bx ⊃Ix) For every x such that x has Being, x is immortal
   B3. (∃x)(¬Cx) There exists some x (Socrates) such that x does not have Becoming.
   B4. ∴ (∃x)(Bx) Therefore, there exists some x (Socrates) such that x has Being (B1, B3, exclusive disjunction)
   B5. ∴ (∃x)(Ix) Therefore, there exists some x (Socrates) such that x is immortal (B2, B4, universal instantiation)

5 One interpretation of Zeno’s paradoxes also lends support to B1, by showing the need for a set of conceptual categories that extends mere Being and not-Being

6 For one enlightening (and amusing) take on begging the question and philosophers who charge their opponents of doing so, see van Inwagen 1983: 18, 101-103.

7 For detailed explication of what the cultural and philosophical expectations vis-à-vis the soul might have been see Brammer 1987.

8 For an account of how representation, and, more broadly, art, might confer epistemic justification, see Young 2000, esp. pp.94-103.

9 I am indebted to David Ciocchi, John Mark Reynolds, Nat Tabris, and a referee from this journal for their contribution to the development of this paper. All errors are, of course, my own.
Andrew Bailey is an undergraduate philosophy major and teaches parliamentary debate at Biola University in La Mirada, California, USA.

E-mail: wrathius@gmail.com