Fictional Objects
Fictional Subjects

Graham Priest

Departments of Philosophy, CUNY Graduate Center and the University of Melbourne

Abstract

In some incontestable sense, fiction involves a relationship with two poles, subject and object: the subject who engages with the fiction, and the objects of the fiction with which they engage. This essay is about the two poles of this relationship. The most notable thing about the object pole is that many of its denizens appear to be non-existent. But is this, indeed, the case? Can there really be non-existent objects? Turning to the subject pole: the subject, unlike many objects of fiction, would certainly appear to exist. But does it? As Buddhism argues, perhaps this is no more existent then the purely fictional objects.

PT Introduction

At a visit to the local bookshop, you pick up a new book. It happens to be the stories of Sherlock Holmes by Arthur Conan Doyle. Over a coffee, you start to read. You read about things that don't really exist: objects such as Holmes and Watson, and states of affairs, such as a cocaine-using detective living in 221B Baker St, London. When you arrive home, you turn on some music to listen to. It is Giacomo Puccini's opera *Madame Butterfly*. Again, you hear of things that don't really exist, such as the unfortunate Cio Cio San, and a visit to Nagasaki by a US naval officer called by the name of B. F. Pinkerton.

Well, to be precise, you read the words on the pages of the book, and you hear the words of the sung libretto. But do these words *really* refer to non-existent objects or states of affairs? Are there really such things?

And what about you?—the subject reading the stories or listening to the opera. Whatever one says about Holmes and Butterfly, this subject would seem to be much more real than either of these. But is this so? Certainly there are visual or auditory experiences taking place. But is there really a subject?—an entity which underlies these experiences, constituting them as a unity? Or is this just as much a mere appearance as a woman committing suicide at the end of a performance of Puccini's opera?

We might summarise these questions in a simple way. In some quite banal sense, fiction generates a relationship between a subject and objects. But what kinds of things, exactly, is that a relationship between?

PT Fictional Objects A Names in Fiction

Let us start with the object pole of the relationship, and get one thing clear straight away. Some names that occur in fictions, such as 'Holmes' and 'Butterfly' refer, if they refer to anything at all, to things that do not seem to exist. But some names that occur in fiction *do* refer to existent entities, such as 'Baker St' and 'Nagasaki'. It might be thought that such names do not refer to the places in question, but to different objects —some fictional Baker St or Nagasaki, quite distinct from the real locations. After all, a detective with the name 'Holmes' lived in Doyle's Baker St, and no such detective ever lived in the real Baker St. So these must be different Baker Streets (since different things are true of them: the principle of the difference of discernibles).

But this is just a confusion. It is certainly true of Baker St that no detective with the name 'Holmes' ever lived in it. But it is not *really* true that such a detective did live in Baker St. What *is* true is that *in Doyle's stories* a certain detective lived in Baker St. So one cannot apply the principle of the difference of discernibles to obtain the conclusion that they are distinct.

Moreover, *both* of these things are true of one and the same Baker St. Otherwise we could not say things like: there is a place in London such that, in Doyle's stories, a certain detective lived *there*, but no such detective actually lived in *it*.[1]

Similarly, when the librettists for Puccini's wrote the text for his opera, the woman they wrote about may be purely fictional,[2] but the Nagasaki they wrote about was the actual Japanese city, intimidated by the US navy's Commander Perry in 1853, and later incinerated by a US atomic bomb. When the librettists wrote the words of the opera, he did not, after all, change the meaning of the word 'Nagasaki': he used it in exactly the same way that anyone else would: to refer to the city in question.

A final example. There is a story entitled 'Sylvan's Box'.[3] It is about an old friend of mine, Richard Sylvan, and is set just after his untimely death. Many of the things in the story are actually true (and not just true-in-the- story). But it is a work of fiction: the central event recounted never really took place. The story is about Richard himself, however. How do I know? Because I wrote the story, and when I used his name in the story, I was referring to him.[4]

So a name employed in a fiction may well refer to an existent object. But of course, many names in fiction do not seem to. Let us say that a name is *fictional* if it is used in some work of fiction. If a fictional name does not seem to refer to an object which exists, let us call it *purely* fictional.

B Reference Failure

Now, purely fictional names either refer or they do not. And if they refer, they refer to something that either exists or does not. So we have three possibilities. Let us consider each of these in turn.[5]

The first is that purely fictional names do not refer at all.[6] Some philosophers, for example, say that, when engaging in acts of fiction, the story-teller, and maybe the story-hearer, merely *pretend* that such names refer.[7] Now, of course, there may well be an act of pretense in fiction. The woman playing Butterfly does pretend to die at the

end of Puccini's opera. But when the actress sings of Nagasaki, she is not pretending to use a name: she *is* using it. In that context, nothing different seems to be going on with her use of 'Pinkerton'. Indeed, the actress may be entirely unaware of whether 'Nagasaki' and 'Pinkerton' are purely fictional names—or even mistaken about matters.

Worse, we now accept that the Homeric stories are mythical: 'Zeus', 'the Minotaur', and their like, are purely fictional. But in the Homeric period of Ancient Greece, the stories were believed to be true. Those who recounted the tales were not *pretending* anything.

Even setting the notion of pretense aside, a greater problem looms. If a purely fictional name does not refer to anything, how are we to account for the truth of some things we say employing the name? We may discount things like 'Holmes lived in Baker St'. This is not really true, as observed. We may even discount things such as 'Holmes does not exist'. Given a certain understanding of reference-failure, we may take 'Holmes exists' simply to be false, and so its negation to be true.[8] These are not the problems.

Problems arise when we consider reports of what happens within works of fiction, such as:

- In Doyle's stories, Holmes lived in Baker St.
- In Greek mythology, Zeus lived on Mt Olympus.

True, the names occurs within the scope of the operator 'In such and such a story/myth', but they are still purely fictional names, and the claims made employing them are true. One might suggest that in such contexts, 'Holmes', for example, refers to an object, one that exists at the worlds that realise Doyle's narratives, but not at the actual world. But then when I make this claim, I am referring to said Holmes. So this is just to say that in my mouth 'Holmes' *does* refer to an object—one that does not actually exist.

Even worse is discourse of a second-order nature, such as:

(*1) Holmes is a purely fictional detective

or 'Doyle created Holmes (in some sense)', or 'Holmes was smarter than Inspector Clouseau' or 'The Homeric Greeks worshipped Zeus'—all of which are plainly true.

One might always invoke the noble art of paraphrase. Thus one might suggest that (*1), for example, means:

(*2) In some work of fiction, Holmes is a detective, but 'Holmes' does not actually refer to anything.

This throws us back to the first kind of context, of course; but now there are new problems. (*1) would seem to entail that:

(*3) *Something* is a fictional detective (namely, the referent of 'Holmes').

But (*2) cannot, since, according to it, 'Holmes' has no referent. So (*1) and (*2)

cannot mean the same thing.

Philosophers being the ingenious creatures they are, they will undoubtedly delight in suggesting other paraphrases. Such jugglings should, of course, be treated on their merits. But it is clear that a desperate rear-guard action is being fought—and a quite unnecessary one if there are other options that work. Are there?

C Existent Objects

This brings us to the second option. Purely fictional names refer to objects, but these are existent objects. Of course, no one really wants to say that Sherlock Holmes or Zeus really existed. So the names must refer to some sort of surrogate. What?

There are a couple of possibilities. One is that the name, for example, 'Holmes' does not refer to Holmes, but to a mental representation of Holmes. Let us call this representation by the name *Holmes**. What, exactly, such a representation is, one might debate; but it is, at any rate, something in the mind (however that relates to the brain).

The most obvious objection to this is that when we say that Holmes does not exist, if we are referring to Holmes* this is just false. We have, then, to reinterpret the notion of existence. Perhaps when we say that Holmes does not exist, we mean something like: he does not exist outside the mind. But this seems all wrong. Suppose I say (truly) that Plato existed but, in contrast, Zeus does not. It would appear that it is *exactly* what I am saying of Plato, that I am denying of Zeus. That is, it would seem to follow that for some property *P*, Plato has the property *P*, and Zeus does not. Clearly, this does not follow from 'Plato exists, but Zeus does not exist outside the mind'. This attributes different properties.

Worse is to come. Suppose I tell a story about some state of the mind. (Maybe in the story I discover it.) Call this state *s*. I wonder whether *s* really exists. I then discover that it does not. In that case I must have been talking about some mental representation of *s*, *s**, and I have just discovered that it does not exist outside the mind. But this is no discovery: I never thought it did.

Or worse again, Holmes was smarter then any actual detective, we may suppose. When we do so, we are not supposing that a mental representation is smarter than any actual detective. Mental representations are not the *kind* of thing that can be smart. Similarly, Nagasaki and the concept of Nagasaki are quite different. When I say that Nagasaki is in Japan, I am talking about the place; I know that I am not talking about a mental concept. Similarly, when I say Butterfly is a tragic woman in a Puccini opera, I am not talking about a concept. Concepts are not the kind of thing that can be tragic women.

What other existent objects might purely fictional names be referring to? Some have suggested that they refer to abstract entities, like numbers—except that, unlike numbers, they can be brought into existence by human creation (maybe like social institutions).[9] The view shares many of the same problems.

First, when we say that Holmes does not exist, we have to reinterpret the notion of existence. This time, one might suggest taking that when we say that Holmes does not exist, we mean that he does not exist as a concrete object. But this seems all wrong. When I say that Plato existed but, in contrast, Zeus does not, it would appear

that it is exactly what I am saying of Plato, that I am denying of Zeus. Thus, it would seem to follow that for some property *P*, Plato has the property *P*, and Zeus does not. This does not follow from 'Plato exists, but Zeus does not exist outside the mind'. These attribute different properties.

Worse: suppose I tell a story about some hypothetical number. (Maybe in this story I discover it.) Call this number *n*. I wonder whether *n* really exists. I then discover that it does not. In that case '*n*' must have been referring to some abstract object, and I have just discovered that it is not a concrete object. But this is no discovery: I never thought it was.[10]

Even worse: if we suppose that Holmes was smarter then any actual detective, we are not saying that an abstract object is smarter than any actual detective. Abstract objects are not the *kind* of thing that can be smart. Similarly, when I say that Nagasaki is in Japan, I am talking about the place; I know that I am not talking about an abstract object. Similarly, when I say that in Puccini's opera Butterfly is a tragic woman, I am not talking about an abstract object. Abstract objects cannot be tragic women. (Take an abstract object, such as the number 3. It makes no sense to say that three is a tragic woman. It's just the wrong kind of thing.)[11]

D Non-Existent Objects

This brings us the the third possibility. Purely fictional names refer to non- existent objects. This is a perfectly common-sense view. We can say that Sherlock Holmes does not exist, and mean this in a perfectly straightforward way. And how it is that we can say other true things about such objects is also clear: the objects simply have the properties being attributed to them. So why not say the obvious?

The answer will itself be obvious to anyone who knows the history of 20th century English-speaking philosophy—though not, perhaps, to anyone else.[12] Non-existent objects had a really bad philosophical press in the 20th century (unlike in the other great periods in Western logic and metaphysics).

In 1948, the American philosopher Willard Quine published a paper, 'On What there is', [13] in which he argued that the only way to express existence is by using the word 'some'. *Some* means *some existent*. To say, then, that some fictional objects do not exist, is a contradiction in terms.

Quine's view seems most implausible. Can't we say—truly—things such as 'Some things don't exist, like Father Christmas, Zeus, Butterfly' or 'I wanted to buy you something for Christmas, but I found out that it doesn't exist' (for example, something owned by Sherlock Holmes, whom I had mistakenly believed to be real). But Quine's paper was enormously influential. Quine's view is now under serious—and quite justified—attack. But it is still a very common one—it is perhaps still the orthodox view: in many places in the English-speaking philosophical world, to say that you think that some things don't exist is taken to be just a shade short of insanity.

In his paper, Quine argues for his thesis by elimination. There is nothing *else* with which one can express existence, so it has to be the quantifier, *some*. Perhaps the most obvious thing that strikes someone who now reads the essay, and who has not taken Philosophy of Language 100—other than its hugely rhetorical content—is this: Quine doesn't even mention, let alone consider, the most obvious candidate to express

existence, namely the verb *exists*, as in 'Nagasaki exists; Lilliput does not'.

This odd omission is perhaps explained by the influence of Bertrand Russell on Quine. In his lectures on logical atomism of 1918, Russell argued that *exists*, construed as a predicate, is meaningless.[14] To say of an object that it exists (or does not exist) is literally nonsense. Existence is a notion that applies only to groups of things, and to say that they exist, is to say that some things are in that group. Russell's view appears to be even more incredible than Quine's.[15] When I was a child I believed that Sherlock Holmes existed. Later on, I learned that he did not. What I learned was nonsense? Russell defends his view with a battery of arguments, which are, frankly, frightful. Yes, I know that that's a very strong claim, and I don't intend to defend it here.[16]

There is a defensive move that may be made here. A name, *n*, is really a covert description, and so is really something of the form: *the thing satisfying condition C.*[16] To say that *n* exists is then to say that something—or, for Russell, some unique thing—satisfies condition *C*. The view that names are really covert descriptions has now been widely discredited, however.[18] The reasons are many. Here is just one. Descriptions display an ambiguity in modal contexts (such as those produced by words like *might* and *must*); names do not. Thus, consider, 'The 44th president of the United States might not have been a man'. This may mean 'Barack Obama might not have been a man'. Or it might mean 'The person who won the election for the 44th presidency might not have been a man'. (Hillary Clinton might have won the election.) These mean quite different things. There is no similar ambiguity in 'Barack Obama might not have been a man'.

Russell's view, to the effect that existence is not a monadic predicate, is often foisted on Kant in some of his remarks in the *Critique of Pure Reason*. The attribution is mistaken, as it must be: *existence* (reality) is one of Kant's categories—distinct, I note, from *particularity* (some) (A80=B106). What Kant actually says in his discussion of the Ontological Argument (A592=B620 to A603=B631), is that existence is not a *determining* (*Bestimmung*) predicate. And what that means is that to say that something is an *X* is the same as to say that it is an existent *X*. The view is itself mistaken. For some sorts of *X*s, existent *X*s are the same as *X*s. Chairs, for example, are things in space and time, and so existent. So all chairs are existent chairs. But this is not true for all *X*s. Some fictional characters (that is, characters that occur in a work of fiction) exist, such as Nagasaki and Sylvan; some do not, such as Lilliput and Pinkerton. So a fictional character need not be an existent fictional character. Be that as it may, Kant's view was not Russell's. For Kant, to say that an object exists is not at all meaningless. Indeed, since the existence of something is a synthetic matter (A598=B626), the discovery that something exists (or does not) can be a very significant matter.

In sum, then, there seems to be no real philosophical bar to accepting the common-sense view that some names denote non-existent objects.[19] The names used in fiction, then, denote. Some denote perfectly existent objects, like Nagasaki and Sylvan; some (the purely fictional names) denote non-existent objects, such as Lilliput and Butterfly. Let us call such objects themselves purely fictional.

E Properties of Non-Existent Objects

What properties, though, do purely fictional objects have? Primarily two kinds. The first comprises those which they possess in virtue of the fiction in which they occur. Thus, it is not true that Holmes has the property of living in Baker St, but it is true that Holmes has the property of (living in Baker St in Doyle's novels). Similarly, Cio Cio San has the property of (committing suicide in Puccini's opera).

To determine what properties something has in a fictional context is not always straightforward. Normally, if the author of a fiction says or shows explicitly that a character is or does something, then that is, indeed, a property of the character in the fiction. But not always. Sometimes it may become clear that a character in a fiction lies or is unreliable in some other way,[20] even the narrator—think merely of Baron Munchausen. Perhaps more importantly, many things may be true in a work of fiction though the author does not say or show so explicitly. Thus, in the Holmes stories, large doses of arsenic kill people, you cannot get from London to Edinburgh in an hour, guns are not made of butter. Doyle says none of these things. It is simply assumed that facts from the real world (or the facts of the world of London *circa* the second half of the 19th century, as Doyle took them to be) are imported—and so may be true *simpliciter*, as well as true in the fiction. What, in general, determines what is, and what is not, imported, is a tough question; but it is not one, fortunately, which we need to tackle here. Intuitively, the notion of truth in a work of fiction is clear enough for present purposes. I note that, in a fiction, an object may have impossible properties. Thus, in 'Sylvan's Box', the box in question—call it *b*—has the property of being empty and occupied by something at the same time.[21] This in no way threatens the Principle of Non-Contradiction. The statements 'in "Sylvan's Box" *b* is empty' and 'in " Sylvan's Box" *b* is not empty' are not contradictories. What *would* contradict the claim ' in "Sylvan's Box" *b* is empty' is the claim 'it is not the case that in "Sylvan's Box" *b* is empty'—a quite different matter. So much for one kind of property of non-existent objects. The other kind comprises those properties that may be attributed, but which do not employ operators of the kind 'In fiction *F* ...'. These are things such as: Holmes is a fictional detective, Holmes does not exist, Holmes was invented (at least in some sense) by Doyle, Holmes is more famous than many real detectives, Holmes is a possible object, Sylvan's box is an impossible object, I am now thinking of Sylvan's box; and so on.

Roughly speaking, the most obvious such properties fall into two kinds. First, there are status properties (*exists*, *is possible*, *is impossible*, and so on.). Secondly there are properties that hold in virtue of the intentional state of some agent directed towards the object (*is being thought about*, *has been heard of by*, *is admired by*, and so on.) Whether there are other sorts of properties may be moot. But one thing we can be sure about is that non-existent objects do not—by definition—have any properties that entail existence. Thus, to exert a gravitation effect on the Moon is to be involved in causal processes, and so existent. Hence, no non-existent object exerts a gravitational effect on the Moon. How, exactly, to determine whether a property is existence-entailing may also be a matter of some dispute. But, again, it is not one we need to go into here. It is time to move to the second half of this essay.

PT Fictional Subjects

A The Sense of Self

So let us consider the subject pole of the relation. Let me start with a word of clarification. In what follows, I will often use the word *I* and its cognates. What I am referring to is the biological organism Graham Priest. (Similarly for, *you*, *we*, and so on, and their cognates.)

When I read a novel or listen to an opera, there certainly seems to be a conscious subject into which the thoughts are entering. We all of us seem to have a sense of self. When one wakes up in the morning after a deep sleep, it is as though a little voice says 'Hello, back again'. Or as Kant put it in more Teutonic terms (B131,132), every mental act is accompanied by the thought *I think*, which constitutes the unity of my thoughts.

So we have a *sense* of self. But do we really have a self? We know that the mind—or the brain whose functioning delivers it—plays tricks. At the back of the eyeball there is a place where the optic nerve joins it. There are no rods or cones there, so the joint produces a blindspot in the field of vision. Normally, though, we are quite unaware of this, since the brain "fills in the visual gap". In a similar way, there is a familiar illusion known as the *Phi Phenomenon* (made use of in the production of movies). Suppose there is a sequence of lights such that from left to right, say, each light flashes momentarily after the one before it. When one looks at this, one actually sees something moving from left to right. The brain "fills in the gaps". Maybe the self of which one has a sense is just the brain filling in the gaps between mental events, as it were, to create the illusion of something that does not really exist.

B Buddhism and Modern Science

That this is so is, in fact, a very ancient view. It is the theory of mind developed over 2000 years ago in Buddhist philosophy. According to this, the mind is nothing more than an aggregate of mental events, causally connected in certain ways to each other and to the body. There is no self over and above—or under and beneath—them. For obvious reasons, the view is called no-self (*anātman*).[22]

The view is also a very modern one, receiving support from developments in cognitive science. Daniel Dennett describes the situation as follows:

> There is no single, definitive "stream of consciousness," because there is no central Headquarters, no Cartesian Theater where "it all comes together" for the perusal of a Central Meaner. Instead of such a single stream (however wide), there are multiple channels in which specialised circuits try, in parallel pandemoniums, to do their various things, creating Multiple Drafts [GP: of a narative of the self] as they go. Most of these fragmentary drafts of "narrative" play short-lived roles in the modulation of current activity but some get promoted to further functional roles, in swift succession, by the activity of a virtual machine in the brain. The seriality of this machine (its "von Neumannesque" character) is not a "hard-wired" design feature, but rather the upshot of a coalition of these

specialists.[23]

This view of the mind, then, has both an ancient pedigree and contemporary scientific credentials.

C Look for Yourself

The current science of the mind is, however, in a rapidly developing state. So perhaps it is wise not to put too much weight on current scientific considerations. Why else might one suppose this view of the mind to be correct? A reason was provided by David Hume (who is often held to have a view about the self akin the the Buddhist view). As he puts it in the *Treatise on Human Nature* (I, IV, 6):

> There are some philosophers who imagine that we are every moment intimately conscious of what we call our Self; that we feel its existence and its continuance in existence; and are certain, beyond the evidence of a demonstration, both of its perfect identity and simplicity. The strongest sensation, the most violent passion, say they, instead of distracting us from this view, only fix it the more intensely, and make us consider their influence on *self* either by their pain or pleasure...
>
> 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 anything but the perception... If anyone, upon serious and unprejudiced reflection, thinks he has a different notion of *himself*, I must confess, I can reason no longer with him. All I can allow him is, that he may be in the right as well as I, and that we are essentially different in this particular. He may, perhaps, perceive something simple and continued, which he calls *himself* ; though I am certain there is no such principle in me.
>
> But setting aside some metaphysician of this kind, I may venture to affirm of the rest of mankind, that they are nothing but a bundle or collection of different perceptions which succeed each other with an inconceivable rapidity and are in perpetual flux and movement…[24]

I'm told by people who practice a certain kind of Buddhist meditation that it is exactly an exercise in simply experiencing the constant arising and ceasing of mental states.

Given his empiricism, Hume inferred from the fact that we cannot perceive the self that there is no such thing—or at least, that we have no reason to suppose that there is. That's too fast. It is true that direct experience may give us no reason to believe in the self. But there are now many things which we take to exist which we cannot perceive, such as electrons and dark matter. What it *does* mean, is that the self has to be considered as a *theoretical posit*, like the scientific entities just mentioned. And to be legitimate, it therefore has to earn its theoretical keep.[25] How?

D Causation and Unity

There is a methodological principle termed *Ockham's Razor*. One should not believe that something exists unless there is good reason to do so. And in the case of all theoretical posits, this means that the posit must have explanatory value. If everything can be explained without it, one should not believe in it. So what might the existence of a self explain that cannot be explained by other things that we are already committed to?

The obvious thought is the Kantian one, that it is the self—whatever that is—which accounts for the unity of experience. Some mental events hang together in a way that others do not. It is precisely the self that is, supposedly, responsible for this.

Let us pass over the somewhat tricky question of how, exactly, the self might turn this trick; but look directly at the question of whether there are other possible explanations for this unity. The unity has both a synchronic aspect and a diachronic aspect. Let us consider each of these in turn.[26]

Synchronic. A motor bike drives past. I see it and hear it. Though one sensation is visual, and the other is auditory, they work together to produce a unitary experience. By contrast: you also see the bike go past, so we both have visual experiences of the bike, but there is no sense in which they are unified in the same way.

This distinction can, however, be explained in simple causal terms. There are causal relations between my auditory and visual sensations which do not hold between your visual sensations and mine. Specifically, the visual and auditory inputs of my brain are processed by different areas of my brain (the visual and auditory cortexes), but these two cortexes communicate with each other in a process of multi-sensory integration to deliver the resulting mental experience. By contrast, there is no similar causal integration between your visual sensation and mine.[27]

Diachronic. This can be past-oriented or future-oriented. Past oriented. Yesterday I saw a road accident. Today I have a visual memory of it. For me, the visual and memory events are integrated, in a way that any of your visual events are not related to my memory. But again, there is a perfectly natural causal explanation of this integration. When I saw the accident, the results from the visual cortex were encoded in the part of the brain responsible for episodic memory (the limbic system). These can be stimulated to generate the visual memory. Obviously there is no similar connection between your visual experience and my memory.

Future oriented. Tonight I have a drink. Because of its pleasant effect, I drink too much. Tomorrow I have a hangover, with its painful mental symptoms. This evening's desire, and tomorrow's headache, go together. However, if you desire to drink, and drink too much, your hangover is not part of my experience. Again, however, there is a perfectly causal explanation of this. I desire to drink, so I drink. The alcohol enters my body, and the overdose gives me a mild case of alcohol poisoning, which my brain monitors the next day, giving rise to the headache. There is no similar causal chain between your drinking, and any headache I might have the next day.

For similar reasons, it makes sense for me not to drink too much tonight if I don't want to have a hangover tomorrow—in a way that it makes no sense for me to try to stop *you* drinking so that *I* don't get a hangover. Thus, the causal relations also make

sense of agency without a self.

E The Illusion of Self

Given all this, it would indeed appear that the experience of self is an illusion: there is no such thing. Of course, illusions can be useful. If you look in a mirror, what is behind you appears to be in front of you. This is an illusion; but it may be a useful one, since it lets you know what is behind you.[28] And one can well imagine that an illusion of self is useful in evolutionary terms (which might, therefore, explain why certain kinds of biological organisms have it). Plausibly, a creature with the illusion of self is more likely to survive and pass on its genes.

However, at least in Buddhist terms, the illusion is pernicious. It generates a spurious attachment at the root of much unhappiness. And once one comes to understand that the object of the attachment is non-existent, it makes no sense to maintain the attachment—any more than it makes any sense to be attached to Butterfly and her well-being, if one knows her to be non-existent. With this disappearance of attachment, the unhappiness it causes will also disappear.

Well, that's the theory. What truth there is in it, is not germane to the present matter.[29] The important point here is that, given that these considerations are correct, the self is an illusion.

PT Conclusion

The fact that the self is an illusion does not mean that it is not an object. In an illusion, we are phenomenologically aware of something, such as the moving object in the case of the Psi Phenomenon. But the object in question is non-existent. That is part of what is involved in the experience being an illusion.

Indeed, the illusory self is a purely fictional object.[30] It is just a character in a fiction that the brain weaves, part of a fictional narrative that the brain fashions, in much the same way that Doyle's brain also fashioned fictional narratives about a non-existent detective.

And of course, this means that the self does not have the properties one may take it to have, such as existence, constancy, unifying power, and so on, any more than Holmes has the property of actually having lived in Baker St or Butterfly has the property of actually having lived in Nagasaki. It is in Doyle's stories that Holmes lives in Baker St, and in Puccini's opera that Butterfly lived in Nagasaki. In the same way, it is in the fiction created by the brain that the self has the properties in question.

But the fact that the self is an illusion tells us more. We are unlikely to suppose that things are as they are said to be in the Doyle stories—unless we mistakenly take the stories to be history rather than fiction. But it is a feature of illusions that we precisely *do* have a tendency to suppose that things are as the illusion shows us—especially if we do not know that matters are illusory; and even when we do, the illusion may be very hard to shake off. We naively take it to be the case that the self exists, is constant, has unifying power, and so on, not realising that it is only in the fiction created by the brain in which these things are true.

And—to return to where we started—fiction delivers a relation which holds

between a subject pole and an object pole. The objects of the object pole may exist; more normally, they do not. The subject pole does not exist at all. It is a non-existent object of a very peculiar kind: an illusory object which, in the illusion, can grasp other non-existent objects. In the fictional narrative created by the mind, it is a purely fictional object which, in the that narrative, can grasp other purely fictional objects.[31]

Notes

1. One might hold that this claim begs the question. But if reference in a work of fiction is *ipso facto* to a different person, then the same is true of reference to: properties (like *detective*), substances (like *cocaine*), numbers (like *three*). These words clearly have the standard meaning (and so reference) when Doyle uses them in his stories—or we could not understand them.

2. Actually, one might contest this. The libretto is based on a story by J. L. Long, which is, in turn, based on a semi-autobiographical novel by P. Loti. This raises the thorny question of the identity conditions of characters across works of fiction. Fortunately, we don't need to go into this here.

3. See Priest 1997.

4. One might reply here: yes, you were referring to Richard, but the name in the text refers to someone else. In reply, one can only ask how the words magically changed their meaning when written on the page. Suppose I write a salacious story about you, and you sue me. The claim that your name did not refer to you simply because it was written down would be laughed out of court.

5. There is a substantial literature on this area. For a survey, see Kroon and Voltolini 2011. It is impossible to do full justice to it in an essay of this kind, so I shall not try. Apologies to anyone whose favourite theory does not get mentioned.

6. Perhaps the earliest version of this view is to be found in Frege 1892.

7. Versions of such a theory can be found in Walton 1993, Recanati 2000, and Kripke 2013: ch. 1.

8. This is the view deployed in so called negative free logics. See Priest 2008: 13.4.

9. Versions of this view are to be found in Thomasson 1999, 2003, and Kripke 2013: ch. 4.

10. Or again, Le Verrier postulated the existence of a sub-Mercurial planet, Vulcan. His theory turned out to be false. It seems somewhat bizarre to say that he postulated the existence of an

abstract object, which, on this account, he did.

11. Some thinkers (e.g., Kripke 2013) combine the view that fictional names do not refer in the discourse of the fiction with the view that they refer to abstract objects in discourses about the fiction. This would seem to be an unstable position. Given it, it is hard to make sense of the following sort of scenario. A tourist to London asks a local policeman, 'Where is the house is Baker St in which Sherlock Holmes lived?' The policeman replies that there is no such place: Holmes was just a fictional character. According to the view in question, the tourist is making a claim that is true in a fiction; so in their mouth 'Sherlock Holmes' has no referent. But the policeman is taking about the work of fiction; so in their mouth, they are referring to an abstract object. Tourist and policeman would then seem to be talking at cross purposes.

12. For a longer discussion of the following matters, see Priest 2016: ch. 18.

13. Quine 1948.

14. See Pears 1972: Lecture 5.

15. Quine is not committed to the view that a monadic existence predicate is meaningless. 'There is y such that $y = x$' is such a predicate, and it is perfectly meaningful. For him, it is just vacuously true of any x.

16. I put my intellectual money where my mouth is in Ch. 18 of the second edition of Priest 2005.

17. This is essentially Russell's 1905 theory of definite descriptions. Russell's theory was indeed appealed to by Quine, to argue that names themselves have no existential import.

18. Largely due to Kripke 1972.

19. For a full defence of the view, see Priest 2005.

20. These are sometimes known as unreliable narrators.

21. If, therefore, one understands the notion of truth in a fiction in terms of what is true at those worlds that realise the fiction, the worlds in question may be impossible worlds. One may suggest that it is not the case that in the fiction the box is both empty and not empty. The narrator is just unreliable. (See Nolan 2007.) Of course, there are such interpretations of the story. The story could, in fact, be interpreted in many different way. It remains the case that there is a natural and straightforward interpretation, according to which, in the fiction, the box is empty and not empty. That is the interpretation I intended, and the one I am talking about here.

22. See Siderits 2007: ch. 3.

23. Dennett 1993: 253-4. The book reviews the evidence and mounts the case for the view. See, especially, Part II of the book.

24. Selby-Bigge 1978: 251-2.

25. The point is perfectly orthodox in Buddhist philosophy of mind. Thus, Vasubandhu (fl. 4th or 5th Century CE), in his discussion of *anātman* notes that if there is reason to believe in a self, it must either be perceived or inferred. See Duerlinger 2003: 73-4.

26. The most sophisticated Buddhist discussion of the matter I know is by Vasubandhu in his 'Refutation of the Theory of Self' (Ch. 9 of his *Abhidharmakośa-Bhāṣya*). See Duerlinger 2003: 71-110. For further discussion, see Carpenter 2014: ch. 6.

27. Of course there are also causal connections between your perception and mine; but they are just of the wrong *kind* to produce the unity in question. Similar points can be made for the examples that follow.

28. Note that an illusion does not have be a delusion. Delusion involves a false belief.

29. See, e.g., Carpenter 2014: chs. 1, 2.

30. Note that I am not suggesting that all purely fictional objects are illusions. For normal adults, the Sherlock Holmes tales do not deliver an illusion.

31. For very helpful comments on earlier drafts of this essay, thanks go to Ridvan Askin, Franz Berto, and Amber Carpenter.

References

Carpenter, A. (2014), *Indian Buddhist Philosophy*, Durham: Acumen.

Dennett, D. (1993), *Consciousness Explained*, London: Penguin Books.

Duerlinger, J. (tr.) (2003), *Indian Buddhist Theories of Persons*, London: Routledge.

Frege, G. (1892), 'Über Sinn und Bedeutung', *Zeitschrift für Philosophie und philosophische Kritik*, 100: 25–50; translated into English as pp. 56-78 of M. Black in P. T. Geach (eds. and trs.), *Translations from the Philosophical Writings of Gottlob Frege*, Oxford: Blackwell, 1980.

Kripke, S. (1972), 'Naming and Necessity', pp. 253-355 of D. Davidson and G. Harman (eds.), *Semantics of Natural Language*, Dordrecht: Rei- del; reprinted as *Naming and Necessity*, Oxford: Basil Blackwell, 1980.

Kripke, S. (2013), *Reference and Existence*, Oxford: Oxford University Press.

Kroon, F., and Voltolini, A. (2011), 'Fiction', in E. Zalta (ed), *Stanford Encyclopedia of Philosophy*,

Nolan, D. (2007), 'A Consistent Reading of "Sylvan's Box"', *Philosophical Quarterly* 57: 667-732.

Pears, D. F. (ed.) (1972), *Russell's Logical Atomism*, London: Fontana.

Priest, G. (1997), 'Sylvan's Box', *Notre Dame Journal of Formal Logic* 38: 573–82; reprinted as section 6.6 of Priest 2005.

Priest, G. (2005), *Towards Non-Being*, Oxford: Oxford University Press; 2nd (extended) edn, 2016.

Priest, G. (2008), *Introduction to Non-Classical Logic*, 2nd (extended) edn, Cambridge: Cambridge University Press.

Quine,W. V. O. (1948), 'On What There Is', *Review of Metaphysics* 48: 21–38; reprinted as ch. 1 of *From a Logical Point of View*, Cambridge, MA: Harvard University Press.

Recanati, F. (2000), *Oratio Obliqua, Oratio Recta*, Cambridge, MA: MIT Press.

Russell, B. (1905), 'On Denoting', *Mind* 14: 479–93; reprinted as ch. 5 D. Lackey (ed), *Essays in Analysis*, London: Allen & Unwin,1973.

Selby-Bigge, L. A. (ed.), (1978), *David Hume: a Treatise on Human Nature*, 2nd edn, Oxford: Oxford University Press.

Siderits, M. (2007), *Buddhism as Philosophy*, Aldershot: Ashgate.

Thomasson, A. (1999), *Fiction and Metaphysics*, Cambridge: Cambridge University Press.

Thomason, A. (2003), ‘Speaking of Fictional Characters’, *Dialectica*, 57: 205-223.

Walton, K. L. (1990), *Mimesis as Make-Believe*, Cambridge, MA: Harvard University Press.