

The Doxastic Requirement of Scientific Explanation and Understanding

SEUNGBAE PARK

Division of General Studies, Ulsan National Institute of Science and Technology,
Ulju-gun, Ulsan 689–798, Republic of Korea
nature@unist.ac.kr

ORIGINAL SCIENTIFIC ARTICLE / RECEIVED: 01–12–13 ACCEPTED: 17–09–14

ABSTRACT: Van Fraassen (1980) and Winther (2009) claim that we can explain phenomena in terms of scientific theories without believing that they are true. I argue that we ought to believe that they are true in order to use them to explain and understand phenomena. A scientific antirealist who believes that scientific theories are merely empirically adequate cannot use them to explain or to understand phenomena. The mere belief that they are empirically adequate produces neither explanation nor understanding of phenomena. Explanation and understanding are the benefits of science only a scientific realist is entitled to.

KEYWORDS: Belief, empirical adequacy, explanation, Moore's paradox, truth.

1. Introduction

In this paper, scientific realism is defined as the view that successful scientific theories are (approximately) true, and scientific antirealism as the view that they are merely empirically adequate. A theory is empirically adequate if and only if whatever it says about observables is true. Thus, a realist believes what a successful theory says about unobservables and observables, whereas an antirealist does not believe what it says about unobservables but believes what it says about observables. The issue I want to explore in this paper is whether the antirealist can use a scientific theory to explain and understand phenomena. I will argue that the belief that an explanatory theory is true is built into any coherent act of explaining and understanding phenomena. Consequently, the antirealist can neither explain nor understand phenomena in terms of scientific theories.

My thesis clashes with what Bas van Fraassen says about acceptance. He claims that “*acceptance of a theory involves as belief only that it is empirically*

adequate" (1980: 12), and that if "the acceptance is at all strong, it is exhibited in the person's assumption of the role of explainer, in his willingness to answer questions *ex cathedra*" (1980: 12). On van Fraassen's account, we can assume the role of explainer with the mere belief that an explanatory theory is empirically adequate. Thus, we can explain phenomena in terms of a scientific theory without believing that it is true. Van Fraassen is not alone on this matter.

An instrumentalist believes that a successful scientific theory is useful but he does not believe that it is true. Rasmus Winther (2009) asserts that the instrumentalist can nonetheless use a scientific theory to explain phenomena:

We are told that there are two attitudes we can take towards a theory: believing that it is literally true, or merely believing that it is instrumentally useful for predicting, explaining, and intervening. The instrumentalist holds the second attitude towards many more theories than the realist. (Winther, 2009: 376)

Note that on Winther's account of instrumentalism, a scientific theory can be instrumentally useful for *explaining*. His account implies that explanation is possible merely on the belief that an explanatory theory is instrumentally useful. Thus, explanation is possible without the belief that an explanatory theory is true.

Contrary to what van Fraassen and Winther maintain, I will argue that we can neither explain nor understand phenomena in terms of scientific theories unless we believe that they are true. In Section 2, I will provide an example to illustrate that Moore's Paradox arises when we attempt to explain and understand phenomena in terms of a scientific theory we do not believe. In Section 3, I will argue that the mere belief that a theory is empirically adequate can generate neither the explanation nor the understanding of phenomena. In Section 4, I will expose a psychological problem and an ethical problem with the antirealist position that we can explain phenomena in terms of a scientific theory without believing that it is true.

2. Moore's Paradox

2.1. *Explainer*

Consider, for example, that two pieces of cold metal become hot when rubbed together at high speed. Why do they become hot? The kinetic theory of heat claims that heat is the mean kinetic energy of molecules, that a material object is composed of molecules, and that the friction between the two pieces of cold metal causes their constituent molecules to be agitated. In short, the two pieces of cold metal become hot because the molecules are agitated. Imagine that the antirealist utters the following two sentences:

- (1) The two pieces of the cold metal get hot because the molecules are agitated.
- (2) But I don't believe that the molecules are agitated.

The antirealist utters (1) to explain the heat phenomenon, and then he utters (2) to be truthful to antirealism. Recall that according to antirealism successful theories are merely empirically adequate. Antirealism prevents him from believing that the kinetic theory of heat is true. So he cannot believe that the molecules are agitated, and he has to say (2).

It sounds odd, however, to utter (1) and (2) conjointly. To assert both (1) and (2) involves Moore's Paradox. George Moore (1993: 207–212) observes that it is absurd to assert a sentence of the form: *p*, but I do not believe that *p*. This sentence form is exemplified by the consecutive utterances of (1) and (2). It is for this reason that I claimed elsewhere (2014a: 12–13) that Moore's Paradox prevents the antirealist from using a scientific theory to explain phenomena. If you want to use a scientific theory to explain phenomena, you should believe that it is true.

I am not claiming here that a scientific theory has to be true in order to yield an adequate explanation. I am fully aware of the historical fact that false theories, such as the Ptolemaic theory, the phlogiston theory, the caloric theory, and the ether theory, were used to explain phenomena. What I contend here is rather that past scientists could not coherently explain phenomena in terms of the past theories without believing that they were true. Thus, what is required for an explanation to be coherent is not the truth of a theory but belief in the truth of the theory. Let me emphasize that my thesis is not refuted by the historical fact that false theories adequately explained phenomena.

2.2. Explainee

So far, I have argued that an explainer must believe that the scientific theory he invokes is true in order to explain phenomena on its terms. How about the explainee? In my view, the explainee must also believe that the theory his explainer invokes is true in order to understand phenomena on its terms. The explainee cannot understand phenomena in terms of the theory if he does not believe that it is true. So I will argue in this section.

Imagine that you observe two pieces of cold metal becoming hot when rubbed at high speed. You wonder why that happens. Your epistemic colleague gives (1) below as an explanation, but you respond to it by uttering (3):

- (1) The two pieces of cold metal are getting hot because the molecules are agitated.
- (3) I now understand, but I don't believe that the molecules are agitated.

You said “I now understand” to express your understanding of the phenomenon, and then you said “but I don’t believe that the molecules are agitated” to express your commitment to antirealism. It sounds odd, however, to utter these two sentences consecutively.

Why is the speech act odd? Consider the sentence ‘I now understand.’ Your utterance of it indicates that you formed a new belief as a result of hearing (1). Specifically, after being exposed to the explanation, you formed the new belief that the molecules are agitated. Thus, when you assert the sentence ‘I now understand,’ you indirectly express your new belief that the molecules are agitated. Therefore, when you said, “I now understand, but I don’t believe that the molecules are agitated,” you indirectly committed Moore’s Paradox.

The preceding example indicates that the belief that molecules are agitated is required to generate the theoretical understanding of why the two pieces of cold metal become hot. To generalize, an explainee can gain an understanding of a phenomenon in terms of a scientific theory, only if he believes that the theory is true. Since the antirealist does not believe that a scientific theory is true, the theoretical understanding of phenomena is beyond his cognitive scope. I am contending not that a scientific theory has to be true to yield an understanding of phenomena, but that we have to believe that a scientific theory is true in order to understand phenomena on its terms.

2.3. Changing the Modes of Speech

In the face of my objection that belief in the truth of a theory is required to explain and to understand phenomena on its terms, the antirealist may change his modes of speech, replacing (1) with (4):

- (1) The two pieces of cold metal are getting hot because the molecules are agitated.
- (4) The kinetic theory of heat says that the two pieces of cold metal get hot because the molecules are agitated.

Note that (4) is not an assertion about the heat phenomenon but an assertion about the kinetic theory of heat. When asserting (4), the antirealist is not himself explaining the phenomenon in terms of the kinetic theory of heat, but is rather reporting an explanation of the phenomenon in terms of the kinetic theory of heat. You can report an explanation without believing that it is true. Moore’s Paradox is not involved in uttering (4), and then adding, “But I don’t believe that the molecules are agitated.” Thus, (4) rescues the antirealist from Moore’s Paradox. So it seems.

The fact remains, however, that the antirealist cannot explain the heat phenomenon using the kinetic theory of heat because he does not believe that it is true. He can at best describe an explanation invoking the kinetic

theory of heat, as (4) illustrates. Furthermore, his explainees may ask him a disconcerting question: “Do you believe what you reported? In other words, do you believe that the two pieces of cold metal get hot because the molecules are agitated?” He cannot answer affirmatively because he is an antirealist. He cannot answer negatively either because if he does, his explainees would discover that he conveyed to them what he did not believe, and hence they would refuse to believe what he reported to them (Park 2014b: 138). They would ask the antirealist a disturbing question: “Why should we believe what you don’t believe?” Since they do not believe what the antirealist reported to them, the original puzzle remains. Why do the two pieces of cold metal get hot when rubbed at high speed?

If (4) relieves explainees of the puzzle at all, that is because they believe that the kinetic theory of heat is true. The explanatory power of (4) derives from the belief that the kinetic theory of heat is true. Suppose that the antirealist gives you (4) as an explanation, but you respond to him by saying (3):

(3) I now understand, but I don’t believe that the molecules are agitated.

It sounds odd to utter (3) in this context as well. The oddity stems from the fact that by saying, “I now understand,” you indirectly say, “The molecules are agitated.” Therefore, you also commit Moore’s Paradox when you say (3) in this context.

So what? It does not matter whether an explainer uses the mode of speech in (1) or the mode of speech in (4). Irrespective of whether the explainer explains the heat phenomenon in terms of the kinetic theory of heat or he simply reports what the kinetic theory of heat says about the heat phenomenon, his explainees form the belief that the molecules are agitated, and this new belief is what produces their understanding of the phenomenon. Again, you have to believe that an explanatory theory is true in order to understand phenomena on its terms.

3. Empirical Adequacy

The antirealist does not believe what a scientific theory says about unobservables, but he believes what it says about observables, i.e., he believes that it is empirically adequate. To use the example of the kinetic theory of heat, the antirealist does not believe that the molecules are agitated, but he believes that any two pieces of cold metal invariably get hot when rubbed at high speed. Within this doxastic framework, the antirealist might give the following explanation of why the two pieces of cold metal get hot:

(5) The two pieces of cold metal get hot because any two pieces of cold metal invariably get hot when rubbed at high speed.

(5) should be tempting to the antirealist because it matches up with his belief that the kinetic theory of heat is merely empirically adequate, and because he does not commit Moore's Paradox, even if he adds that he does not believe that molecules are agitated after uttering (5).

Two comments about (5) are in order. First, (5) is not a theoretical explanation but an observational explanation. In a theoretical explanation, an observable event is explained in terms of at least one theoretical event. (1) is an example of a theoretical explanation because it is a theoretical event that molecules are agitated. In an observational explanation, an observable event is explained exclusively in terms of observable events. Therefore, even if (5) is a tenable explanation, it is only an observational explanation, and the thesis stands that a theoretical explanation is beyond the antirealist's reach.

Second, it is not clear whether (5) is a tenable observational explanation or not. Of course, some observational explanations are legitimate. For example, the traffic accident occurred yesterday because the driver was drunk. (5), however, is not similar to such a legitimate observational explanation. As Alan Musgrave (1988: 242) points out, it is rather analogous to the dubious observational explanation that a swan is white because all swans are white. It is absurd in general to explain an event in terms of another event temporally posterior to it. Imagine that the first swan was born in the distant past, and that it was white. Why was the first swan white? To say that it was white because all swans are white amounts to saying that it was white because all the swans to be born later would be white. Such an explanation is absurd.

Similarly, imagine that Count Rumford, the proponent of the kinetic theory of heat, rubbed two pieces of cold metal for the first time ever. He was puzzled over why they got hot. The antirealist came along and said, "They got hot because any two pieces of cold metal invariably get hot when rubbed at high speed." To say so is to explain an event in terms of the events temporally posterior to it because by hypothesis no one rubbed two pieces of cold metal before Rumford. For this reason, belief in the empirical adequacy of the kinetic theory of heat cannot generate an adequate explanation of the first occurrence of the heat phenomenon.

The problem of explaining the first event in a domain should not be underestimated. There were many times in the history of science when scientists discovered new phenomena by using then the state-of-the-art instruments, such as electron microscopes and particle accelerators. If they believed that their theories are merely empirically adequate, they could not explain the phenomena because the phenomena occurred for the first time with the help of the state-of-the-art instruments.

Jerrett Leplin (1997) goes further, arguing that a generalization does not explain its instance regardless of whether the instance is the first event or not

because the generalization needs to be explained in terms of theoretical entities. For example, the generalization that all swans are white does not explain why a swan is white because the color of all the swans needs to be explained in terms of an underlying structure or a deeper principle:

Generalizations do not explain their instances. On the contrary, the apparent satisfaction of a generalization is more remarkable, and more an inducement to seek an underlying structure or a deeper principle, than the individual case. (Leplin, 1997: 23)

Thus, we need to explain why any two pieces of cold metal invariably get hot when rubbed at high speed by appealing to the underlying structure of the pieces of cold metal. We can do this by invoking the kinetic theory of heat.

Leplin's argument, however, is vulnerable to van Fraassen's critical response that a generalization is a brute fact, i.e., it does not cry out for an explanation. We do not need to explain a generalization in terms of theoretical entities:

So here the anti-realist must similarly say: that the observable phenomena exhibit these regularities, because of which they fit the theory, is merely a brute fact, and may or may not have an explanation in terms of unobservable facts 'behind the phenomena' – it really does not matter to the goodness of the theory, nor to our understanding of the world. (van Fraassen 1980: 24)

On van Fraassen's account, it is a brute fact that any two pieces of cold metal invariably get hot when rubbed at high speed. There might be a theoretical explanation of the generalization. Such an explanation, however, is not necessary for our understanding of the world.

It is not clear, however, why a generalization does not cry out for an explanation when an instance does. After all, the generalization is nothing but a collection of instances similar to the original instance which cried out for an explanation. If the color of a swan is a mystery, so is the color of all the swans. To say that a swan is white because all swans are white is merely to replace a mystery with the set of all the mysteries similar to it. You cannot solve a mystery by providing a set of all the mysteries similar to it. It is not the case that if all the similar mysteries are taken together, they cancel out each other, and that as a result the original mystery evaporates. On the contrary, all the mysteries considered together will only intensify the original mystery. For this reason, (5) is a vacuous explanation.

Relatedly, it is circular to explain an instance in terms of its generalization. To simplify the matter, imagine that there are only two swans in the world, and that they are both white. The color of each swan is explained as follows:

- (E1) Swan₁ is white because swan₁ and swan₂ are white.
- (E2) Swan₂ is white because swan₁ and swan₂ are white.

Note that (E1) and (E2) are saying that swan₁ is white because swan₂ is white, and that swan₂ is white because swan₁ is white. Such explanations are circular, casting no light on why swan₁ is white and why swan₂ is white. It is for this reason, I believe, that Musgrave (1988), Leplin (1997: 23), and Park (2014a: 14) claim that a generalization does not explain its instance. It is vacuous to say that some swans are white because all swans are white.

The problem of circularity does not arise, if the explanans differs radically from the explanandum. You can solve a mystery by referring to another mystery, if they are different in kind. For example, it is justifiable to explain the color of a swan in terms of the photons of certain wave lengths bounced off from the swan. Of course, it is a mystery why the photons of the wave-lengths produce the white color. But it is not circular to provide such a mystery to solve the original mystery.

We have to distinguish between being relieved of a puzzle in science and being relieved of a worry in ethics. Imagine that you are invited to a dinner in a foreign country. To your embarrassment, you belch after the dinner. You worry about your behavior. To your relief, however, your host tells you that in that country *all* people belch after a meal. As a result, you are relieved of the worry. The point of this story is that you no longer worry about your behavior, once you are told that all behave as you do. The word 'all' has such psychological power. The power is erroneously transferred from ethics to science. You are puzzled over why a swan is white. Once you are told that all swans are white, you tend to stop puzzling over the color of the swan. You are tempted to think that the color of the swan is explained, and that it is a brute fact that all swans are white. On close examination, however, the color of the swan is not explained at all, and it is not a brute fact that all swans are white.

4. As-If-Realist

The antirealist may argue that he can explain something merely as if he is a realist. He does not believe that the molecules are agitated, but he speaks like the realist, "The two pieces of cold metal are getting hot because the molecules are agitated." When asked whether he believes that the molecules are agitated, he also speaks like the realist, "Yes, I believe that the molecules are agitated." Thus, the antirealist is verbally indistinguishable from the realist. Such an antirealist might be called a doxastic antirealist but verbal realist. He is an antirealist from a doxastic point of view, but a realist from a verbal point of view. He does not commit Moore's Paradox because he does not say, "I don't believe the molecules are agitated."

Is it plausible that we can be doxastic antirealists but verbal realists? I earlier argued that belief is a prior condition for enlightenment, i.e., we have to

believe that an explanatory theory is true in order to understand phenomena on its terms. In my view, when the antirealist explains phenomena merely as if he is a realist, he does not understand why the phenomena occur. He has no idea why the two pieces of cold metal are getting hot, and yet he says that they are getting hot because the molecules are agitated. His explainees may ask him a perplexing question: “Do you understand why the two pieces of cold metal are getting hot?” Of course, he would say yes because he is a verbal realist. But his answer is not an honest one. Thus, the antirealist merely exchanged Moore’s Paradox with dishonesty.

In addition, the antirealist suffers from what psychologists call cognitive dissonance. Leon Festinger and James Carlsmith (1959) argue that if we say something contrary to what we believe, we experience a kind of psychological discomfort called cognitive dissonance. For example, after performing tedious tasks, we form the belief that the tasks were tedious. Suppose, however, that we are offered a monetary reward on condition we answer affirmatively to the question: *Were the tasks interesting?* We say yes and go through cognitive dissonance. We even tend to replace the previous belief with the new belief that the tasks were interesting, thereby returning to cognitive consonance:

If a person is induced to do or say something which is contrary to his private opinion, there will be a tendency for him to change his opinion so as to bring it into correspondence with what he has done or said. (Festinger and Carlsmith 1959: 209)

Festinger and Carlsmith’s psychological finding has an interesting implication on antirealism. The antirealist experiences cognitive dissonance when he explains phenomena merely as if he is a realist. His cognitive dissonance culminates when he answers the question: “Do you understand why the phenomena occur?” Furthermore, he is likely to become a realist as he speaks like a realist. Therefore, it may be logically possible but psychologically unrealistic that we can be doxastic antirealists but verbal realists.

In addition to the psychological problem, there is an ethical problem with explaining something without believing that an explanatory theory is true. The antirealist does not believe that the molecules are agitated, but he says, “The two pieces of cold metal are getting hot because the molecules are agitated.” His innocent explainees would take his explanation literally, thinking that he believes that the kinetic theory of heat is true. As a result, they would believe that the kinetic theory of heat is true, and they would believe more than what the antirealist believes. Thus, the antirealist’s speech act is epistemically irresponsible with respect to his unsuspecting explainees.

I offered an example elsewhere (2014b: 137) to show that it is unethical to explain phenomena in terms of a theory without believing that it is true. Imagine that you are a member of a cult, and that you wonder why complex

things exist in the world. Your cult leader says, "Only an intelligent designer can create the complex things. My god is infinitely intelligent. He created all the complex things in the world." Persuaded of his explanation, you become his fervent follower and donate all of your money to him. It turns out, however, that he did not believe that the god exists. He merely spoke as if he believed that the god exists. From a verbal point of view, he was a theist, but from a doxastic point of view, he was an atheist! In such a situation, you would think that his speech act was unethical, and that he should not have talked as if he believed that the god exists.

What if the antirealist's explainees know that the antirealist speaks merely as if he is a realist? They may treat him in the way he treated them. When he says that the two pieces of cold metal are getting hot because the molecules are agitated, they refuse to believe that the molecules are agitated on the grounds that he himself does not believe that the molecules are agitated. So they still do not understand why the two pieces of cold metal are getting hot. But they say to the explainer that they now understand why the two pieces of cold metal are getting hot. As they say so, the antirealist feels rewarded, thinking that he enlightened them on the phenomena. In reality, however, they still do not understand why the phenomena occur. They are speaking merely as if they are enlightened on the phenomena. Their language misleads the antirealist exactly as the antirealist's language misleads unsuspecting explainees. In short, the antirealist cannot convince anyone of the truth of his explanation and hence he cannot enlighten anyone, if his epistemic colleagues treat him in the way he treats them. In addition, he will be deceived by his epistemic colleagues in return for deceiving them.

5. Implication

So far I argued that it is problematic to explain phenomena in terms of a theory without believing that it is true. How is this thesis related to the celebrated theories of scientific explanation in the literature? Roughly speaking, Carl Hempel's deductive-nomological model (1966: Chapter 5) holds that to explain an explanandum is to deduce it from explanantia which include a law of nature. Michel Friedman's (1974) and Philip Kitcher's (1981) unificatory theory maintains that to explain phenomena is to unify them. Wesley Salmon's causal theory (1978) asserts that to explain phenomena essentially involves identifying causes. Bas van Fraassen's contextual theory (1980: Chapter 5) claims that an explanation is an answer to a why-question, and that the appropriateness of the answer depends on context. None of these theories addresses the doxastic aspect of explanation, although belief in the truth of explanantia is an important condition for an explanation to be adequate from psychological and ethical points of view.

6. Conclusion

Moore's Paradox arises when we explain and understand phenomena in terms of scientific theories without believing that they are true. The antirealist does not believe that a scientific theory is true. So he can neither explain nor understand puzzling phenomena around us in terms of scientific theories. In general, it is circular to explain an instance in terms of its generalization. So the mere belief that a theory is empirically adequate can generate neither explanation nor understanding of phenomena in its domain. The theoretical explanations and understandings of phenomena are the benefits of science only the realist is entitled to.

The antirealist may contend that he can speak merely as if he is a realist. I replied that the antirealist cannot understand himself why an explanandum occurs, that it is psychologically unrealistic that the antirealist can speak merely as if he is a realist, that it is unethical to explain phenomena in terms of a theory without believing that it is true, and that the antirealist cannot enlighten his explainees on phenomena if they treat him in the way he treats them. Antirealism is not an epistemic policy recommendable to those who are interested in explaining or understanding phenomena in terms of scientific theories.

Finally, I must emphasize again that the main thesis of this paper is refuted not when it is shown that a false theory adequately explains phenomena but when it is shown that we can coherently explain phenomena in terms of a theory without believing that it is true. It is a misguided criticism that many false theories adequately explained phenomena in the history of science. The past theories do not count as counterexamples to the main thesis of this paper.¹

References

- Festinger, Leon, & James Carlsmith. 1959. "Cognitive Consequences of Forced Compliance", *Journal of Abnormal and Social Psychology* 58, 203–210.
- Friedman, Michael. 1974. "Explanation and Scientific Understanding", *Journal of Philosophy* 71, 5–19.
- Hempel, Carl G. 1966. *Philosophy of Natural Science* (Englewood Cliffs, NJ: Prentice-Hall).
- Kitcher, Philip. 1981. "Explanatory Unification", *Philosophy of Science* 48, 507–531.
- Leplin, Jarrett. 1997. *A Novel Defense of Scientific Realism* (New York: Oxford University Press).

¹ I am grateful for useful comments to the anonymous referees of this journal.

- Moore, George. 1993. "Moore's Paradox", in *G.E. Moore: Selected Writings*. Ed. by Thomas Baldwin (London: Routledge).
- Musgrave, Alan. 1988. "The Ultimate Argument for Scientific Realism", in R. Nola (ed.), *Relativism and Realism in Science* (Dordrecht: Kluwer Academic Publishers), 229–252.
- Park, Seungbae. 2014a. "A Pessimistic Induction against Scientific Antirealism", *Organon F* 21, 3–21.
- . 2014b. "Accepting Our Best Scientific Theories", *Pensée* 76, 131–139.
- Salmon, Wesley. 1978. "Why Ask, 'Why?': An Inquiry Concerning Scientific Explanation", *Proceedings and Addresses of the American Philosophical Association* 51, 683–705.
- van Fraassen, Bas. 1980. *The Scientific Image* (Oxford: Oxford University Press).
- Winther, Rasmus G. 2009. "A Dialogue", *Metascience* 18, 370–379.