

Topic-neutrality and the Identity Theory

By

Marvin C. Sterling

In this paper I shall try to set forth a coherent formulation of the psychophysical identity theory. I say "coherent formulation," because I believe the version of the theory that I shall advance to be immune to certain difficulties which it encounters in its more usual forms. The position for which I shall argue is reached by *reversing* the main assumptions of the "two-stage exposition" once advocated by J.J.C. Smart. Moreover, I will suggest that by effecting such a "reversal" one is able to avoid not only certain logical objections sometimes raised, but also the counter-intuitive consequences of the "disappearance form" of the identity theory. My chief contentions in the paper will be these: (a) Scientific reports about *brain-events* are topic-neutral; and (b) That brain-events are nothing but sensations is a reasonable hypothesis. (For convenience, I am using the word "sensation" in this paper in a very general sense to cover *all* phenomenological contents or processes.)

I borrow the expression "topic-neutral" from Smart; and I wish to begin by stipulating the exact sense in which I shall be using that expression. I will employ the word "topic-neutral" to describe a certain kind of statement. The following are examples of this kind of statement:

- (1) Something is in the box.
- (2) Someone is knocking at the door.

The key words to notice in the above statements are "something" and "someone." For it is the use of these words that makes the listed statements topic-neutral. Thus, the distinctive feature of topic-neutral statements is their "openness"; such statements do not specify the characteristics of that about which the assertion is made.

Now, there are some circumstances in which one is *limited* to making topic-neutral statements, and other circumstances in which one has the option to make either topic-neutral or "specific" statements. In suggesting that in some situations one is *limited* to making topic-neutral statements I just mean that in those situations the only statements that one *is justified in making* about a particular subject of discourse are topic-neutral. For example, suppose that I hear a knock at my door, but that I cannot see who it is because the door is shut. Assume moreover that I have no reason to *suspect* that it is anyone in particular. Under these circumstances I would not be justified in saying that some specific individual with such-and-such characteristics was knocking; but I would be justified in asserting

that *someone* was knocking, and moreover, I would be *limited* to the use of such topic-neutral language.

Under certain other circumstances, however, I would not be limited in this way, but could make either sort of statement. For suppose that I hear a knock at my door, and that I *can* see who it is, because the door—being made of glass—is transparent. Given this situation, I certainly *could* make a topic-neutral statement to the effect that *someone* is knocking. But I would not be limited to making this sort of statement, for I would be able just as easily to make a statement in which I indicated the observable characteristics of the individual who was knocking; or—in the event it was not a stranger who was knocking—I could appeal to the same characteristics in making a decision to apply a definite name (e.g., “Natalie Cole is knocking at my door”).

In this paper I am interested primarily in situations of the first sort I described above. That is, I shall be concerned with the kind of situation in which one is *limited* to making topic-neutral statements about the subject of one’s discourse. And I would like to make two distinguishable claims concerning this type of situation.

First, in situations where I am limited to the use of topic-neutral language, the characteristics of that about which I make my assertion (i.e., the characteristics of the subject of my discourse) are largely concealed from me. Of course, I must have some minimal information concerning the “something” or “someone” about which I make my claim, for otherwise there would be no basis for making even a topic-neutral statement. Nonetheless, in situations of this sort, I am *for the most part* ignorant about the subject of my topic-neutral discourse. For example, imagine the following situation. I receive a package in the mail; it is, let us say, a small box and is completely opaque, so that I cannot see inside. I remove it from my mail-box, and I notice that it feels hollow. I then shake it a bit, and hear a noise that seems to come from inside. I immediately infer that this noise is caused by the movement of some item inside the box, and I say to myself “Something is in the box.” Moreover, as long as I do not open the box, and have no reason to *suspect* the presence of some particular object, I am limited to the use of such topic-neutral language.

Now, what I wish to emphasize about this is that, in the situation described, the characteristics of the “something” about which I make my assertion are largely concealed from me. Nevertheless, I do have some information about it. I am able to assume, for instance, that it moves around inside and makes a noise when I shake the box; this, in fact, constitutes the main evidence for my topic-neutral statement “Something is in the box.” However—and this is my chief point—I would still be

ignorant *for the most part* about the nature of the “something” that was in the box; it could turn out to be a ring, a cuff-link, a watch, a stone, or any number of other items. And I wish to suggest that the sort of *ignorance* that obtains in this situation is also present in all situations in which one is *limited* to the use of topic-neutral language. So, the first claim that I desire to make is that in every situation of this sort, there naturally arises the question “What exactly is the nature of this *something*?” (Of course, when I say that this question *naturally arises*, I do not mean that everyone in a situation of this sort *will in fact ask* such a question, but only that, given the limitations of this kind of situation, such a question could legitimately be asked.)

Secondly, I submit that in every situation in which one is limited to making topic-neutral statements about a given subject of discourse, *there is the need for a hypothesis*. (I use the word “hypothesis” here in a broad, non-technical sense.) Specifically, a hypothesis about the nature of the “something” in question is needed. Moreover, the principal reason why a hypothesis is needed is that the characteristics of that about which the topic-neutral claim is made are “hidden,” or “concealed” from the speaker.

Imagine once again the above example of a package received in the mail. As long as I refrain from opening the box, and have no reason to *suspect* that it contains some particular object, I am limited to making the topic-neutral statement that *something* is in the box. The question then arises “What exactly is this something that is in the box?” And there is now the need for a hypothesis in reply to this question. Moreover, feeling this need, one might attempt to gather additional evidence on which to base a definite hypothesis. Thus, as I search my memory, it might occur to me that several weeks earlier I had placed a mail-order for a gold ring. Such a memory would constitute evidence for a particular hypothesis, and I might conclude “The something in the box is a gold ring.” I still could not be absolutely *sure* that there was a gold ring inside, but I would now have some grounds for *suspecting* that there was a gold ring inside. And of course, I would now no longer be limited to making topic-neutral statements, albeit I initially was. So, to repeat what I said above, in the sort of situation with which I am here concerned, there is invariably the need for a hypothesis.

Now, in this paper I want to make two major contentions, and the first one is this: Scientists are limited—at least initially—to the use of topic-neutral language in reporting the occurrence of brain-events. Moreover, in keeping with what I said about topic-neutrality above, I will presently call

attention to a couple of consequences that follow from this contention. But before I do that, I would like to bring out explicitly some assumptions that I am making; for the plausibility of the above contention depends on the correctness of these assumptions.

I assume the following two propositions to be true about sense-perception in general: (1) In sense-perception, we are never *directly* conscious of external things and occurrences; and (2) In sense-perception, what we are directly conscious of invariably are nothing but "mental events"—such mental events being causally produced in our consciousness by external things and occurrences. The theory that is embodied by these assumptions is now perhaps less fashionable among philosophers than among scientists, but it is, nonetheless, one with which any student of the history of philosophy is likely to be familiar. And I will use the expression "causal theory" to designate this view (i.e., the conjunction of (1) and (2) above), although I recognize that "causal theory" has not always been employed in exactly this way.

Now, I want to emphasize that in this paper I do not intend to *argue* for the causal theory; I proceed on the assumption that it is correct. However, I would like to clarify something before continuing. A key notion in propositions (1) and (2) is that of being "directly conscious" of something. And as I am using the expression, to say that one is *directly* conscious of something, *x*, amounts to saying that it is precisely *x* *itself* of which one is conscious, and not merely something, *y*, from which *x* may be *inferred*. To further elucidate this idea, let us contrast it with the notion of being "indirectly conscious" of something. I am *indirectly* conscious of something, *x*, just in case I am conscious of something, *y*, from which *x* may be inferred, but am not conscious of *x* itself. (I stress that I am *stipulating* these definitions, and do not assume any general usage.)

For example, consider again the case previously mentioned of a package received in the mail. When I shake the box, I can hear a sound coming from inside; and using this sound as evidence, I may infer that something is in the box. In other words, I am *directly* conscious of the sound produced when I shake the box, while I am only *indirectly* conscious of the something inside producing this sound. Accordingly, what the causal theory asserts is that we are never directly conscious of external things and occurrences *themselves*, but rather, only of certain mental events that (a) are causally produced by external things and occurrences, and (b) constitute evidence from which external things and occurrences may be inferred.

Now, the brain of another person is an example of an external thing, and an event in another person's brain is an example of an external occurrence. More precisely, it is true for each person that the brain and

brain-events of all *other* persons are external. Thus a consequence of the causal theory is that I am never directly conscious of events in another person's brain—and in general, that no one is ever directly conscious of *another* person's brain-events. We could express this consequence of the causal theory in quasi-Kantian language. Accordingly, we could say that the "brain-in-itself" of another person cannot be identified with the visual "appearances" that one experiences when one is looking at another person's brain; we could say further that the *events* that happen to the "brain-in-itself" of another person cannot be identified either with actual or with potentially obtainable visual "appearances." (Thus, for example, the causal theory would rule out the possibility that the future development of more powerful microscopes might eventually make the "in-itself" of brain-events directly apprehensible. The development of more effective magnification techniques would only allow us to obtain additional visual "appearances," on the causal theory.)

Recall again the earlier example of a package received in the mail. Specifically, suppose that I am shaking the box, and that I am listening to the resultant sound made by the item inside. What I wish to stress about this situation is that—if the causal theory is correct—there are definite similarities between how I must talk about this item in the box, and how scientists must talk, at least initially, about the brains of other persons. Let me try to clarify what I am getting at in suggesting this similarity.

We may focus first on the case of the item in the box. And the important thing to note here is that the characteristics of the item in the box are largely concealed from me; that is, I am for the most part ignorant about this "something" that is in the box. However, I *do* have a basis for making certain topic-neutral statements in this situation; for, the sound that I hear when I shake the box constitutes evidence from which I may infer both that *something* is in the box, and that something *is happening* to the something that is in the box. Moreover, because of the "hidden" status of this item in the box, I am limited (initially at least) to the use of such topic-neutral language in talking about it. Furthermore, in any situation in which I am limited to the use of topic-neutral language, I may ask the question "What exactly is this *something* that is the subject of my topic-neutral discourse?" There is also the need for a hypothesis; and once appropriate evidence has been obtained, I am justified in embracing a definite hypothesis about the contents of the box. And the idea, of course, is that there must be *something* whose existence and activities *are the cause* of the sound that I hear when I shake the box.

Similar considerations apply to the case of scientific talk about brains and brain-events. For if the causal theory is true, then the brains of other persons (i.e., the *in-itself* of other persons' brains) are concealed from the scientist. In other words, a consequence of the causal theory is that the scientist is never directly conscious of the brains of other persons as such, but only of certain experiences of his own that are causally produced by the brains of other persons. However, in this case, like that of the preceding paragraph, there *is* a basis for making certain topic-neutral statements. Specifically, certain evidence is obtainable from which scientists may infer both the existence of other persons' brains, and the occurrence of brain-events in other persons. And the evidence for such inferences is got in sense-perception.

For example, the visual "appearances" experienced by the neurosurgeon during surgery are evidence from which the existence of another person's brain may be inferred. And such things as electroencephalograph-readings are evidence from which the occurrence of brain-events in another person may be inferred. Here again, the idea is that there must be *something* whose existence and activities *are the cause* of the effects produced (i.e., the cause of the visual "appearances," and of the EEG-readings). Moreover, because of the "hidden" status of other persons' brains and brain-events, scientific talk about the brains and brain-events of other persons can be translated into explicitly topic-neutral language. For (assuming the causal theory is correct), when I utter the words "brain" and "brain-event," I do not thereby commit myself as regards the *intrinsic* nature of the entities and processes that I refer to, but rather leave open the possibility of their being of *any sort whatsoever*. Such utterances *do* specify certain *extrinsic* properties of their referents; for, when I say "brain," for instance, I do thereby commit myself to the supposition that the "something" to which I refer causally *produces in my consciousness "appearances" of a specific sort*. Nevertheless, apart from their specification of such *extrinsic* properties, the expressions "brain" and "brain-event" are non-committal—if the causal theory is true—as regards the nature of what they refer to.

And to repeat my earlier claim, in every situation in which one is limited to making topic-neutral statements about the subject of one's discourse, at least two things are true: (a) One may ask the question "What is this *something* about which I speak?" and (b) There is the need for a hypothesis in answer to the question "What is this *something* about which I speak?" Now the causal theory implies that we are limited (at least initially) to making topic-neutral statements about brain-events. So, if the causal theory is true, the scientist may ask: What exactly are these

brain-events about which I speak? Moreover, he will be in need of a hypothesis about the nature of brain-events—a hypothesis stating exactly what sorts of things brain-events are. And this brings me to the second major claim that I want to make in this paper, namely: The sentence "Brain-events are nothing but sensations" expresses a reasonable hypothesis.

Let me once again compare this situation with my previous example of a package received in the mail. Remember that I cannot see what is inside of the box, but that I *am* able to get some evidence from which to infer that *something* is in the box (viz., the sound that I hear when I shake the box). Because of the limitations of this situations, there is a question about what exactly is in the box, and I need a hypothesis to answer this question. Furthermore, I am able to obtain additional information on the basis of which I suggest the hypothesis "The something in the box is a gold ring." Now, what I would like to stress here is that the sentence "Brain-events are nothing but sensations" in the brain-events case serves the same function as the sentence "The something in the box is a gold ring" in the case of the package received in the mail. I shall now attempt briefly to say why I believe that the sentence "Brain-events are nothing but sensations" is a reasonable hypothesis.

It seems to me that the main evidence for the hypothesis "Brain-events are nothing but sensations" is provided by the *verbal reports* of patients who undergo a particular kind of operation. What I have in mind is the sort of surgical procedure for which the Canadian neurosurgeon Wilder Penfield is famous. Let me try to describe what I am getting at. Imagine that a neurosurgeon is performing brain-surgery. And suppose that during the course of this operation, he electrically stimulates a point on the patient's neo-cortex. From the fact of his having administered this electrical stimulation, the neurosurgeon may infer that a brain-event is occurring; and we may assume that the sentence "A brain-event is occurring" is translatable into topic-neutral language. In other words, at this point, the neurosurgeon can infer that *something* is going on, but he is not as yet in a position to say exactly what this something is. (For the causal theory implies that our neurosurgeon is not directly acquainted either with the patient's "brain-in-itself," or with events that happen to the patient's "brain-in-itself.")

But now, suppose that immediately subsequent to the electrical stimulation of his brain, the patient says "I am having a sensation of kind X." Furthermore, suppose that all patients whose brains are stimulated in this way consistently report the having of sensations (or of sequences of sensations) immediately upon the administration of such stimulation; that is, let

us assume that this correlation between the production of brain-events by a surgeon, and the report of sensations by his patient is repeatedly borne out. Would this not constitute strong evidence in support of the hypothesis that brain-events are nothing but sensations? I submit that it would. And, as a matter of fact, such a correlation between the production of brain-events and the report of sensations has been repeatedly demonstrated both by Wilder Penfield and by later investigators. This then is why I am suggesting that "Brain-events are nothing but sensations" is a *reasonable* hypothesis.

A non-physicalist monism is the sort of metaphysical theory that seems to be indicated by the foregoing observations. Moreover, two important advantages of this version of the identity theory are: (1) Unlike some physicalist versions of the identity theory, it does not encounter the difficulty of a *difference* or *incompatibility* between physical and mental properties—since just *one* set of properties is affirmed; and (2) Unlike the "disappearance form" of the identity theory, it is able to avoid the "property objection" without having to rely on the highly implausible assertion that mental events—which we immediately experience—do not exist.

University of Oklahoma