

Empirical Research Cannot Rescue the Disappearance Form of the Mind-Body Identity Thesis

By Dr. Greg Bahnsen[\[1\]](#)

A Reprieve for Materialism?

Realizing that some formulations of the thesis that mental processes are identical with brain processes encounter trenchant objections, Richard Rorty proposes a new statement of the materialist position in his article, "Mind-Body Identity, Privacy, and Categories."[\[2\]](#)

Rorty portrays the identity theorist as impaled on the horns of a dilemma: either he holds to strict identity, in which case mental processes and brain processes have all predicates in common, or he does not. If he does, then he ends up making category mistakes and using meaningless expressions - e.g., "This brain process is false," or "This thought is located six feet above the ground." But if he does not, then he has abandoned materialism for a mere correlation between mental and brain processes.

J. J. C. Smart suggests that category mistakes can be avoided by translating sentences which seem to express them into topic-neutral language, but Rorty finds this unpromising. Such a ploy would always be open to the demand to provide a suitable translation, and then open to the criticism that the proffered translation was inadequate. The materialist would need to "neutralize" the original assertion and then defend the adequacy of this translation by appealing to criteria for translation which do not, somehow, approach the position of mere correlation again.

Nevertheless Rorty is not ready to abandon materialism, believing that there is a better way to express or formulate the position. He calls his view the "disappearance form of the identity thesis" [hereafter DFIT]. In his assessment, DFIT has the advantage of avoiding category mistakes, going beyond mere correlation, and avoiding disputes over adequacy of translation. DFIT would thus be a strong expression of materialism, but not one subject to the common linguistic objections of the non-materialist philosophers. The virtue of DFIT would be that it is philosophically plausible and, as such, subject only to empirical confirmation or disconfirmation.

Rorty makes it clear that he aims simply to show that DFIT "makes sense," not that any prediction of its empirical confirmation is true. That is, since DFIT is a proleptic claim, Rorty wishes to maintain merely that DFIT *could* be true. He denies that current ways of speaking or any

classification of linguistic expressions can block the results of empirical inquiry. So then, if DFIT is to be discounted, it cannot be by means of philosophical considerations, but only through the use of empirical methods. Rorty simply contends that DFIT is a sensible position to affirm, one whose truth-status must be decided later.

Rorty's claim is important because it would salvage the materialist position for at least a time. If Rorty is right, materialism still has a legitimate claim on our metaphysical attention and reflections. It cannot be dismissed without the verdict of the scientists, and the "politically correct" scientific community of our day is strongly disinclined to dismiss a materialist view of man and the world. Rorty wants us to believe that materialism is a viable position which must be taken seriously. Non-scientists must not presume to reject materialism on philosophical grounds. The implications of materialism for epistemology, ethics, anthropology, religion (to mention but a few) are notorious, and so Rorty's claim is worth examining. Does materialism deserve a stay of life?

Let's ask, then, whether DFIT can be philosophically discounted in advance of empirical research. Unless I am missing something, I believe that it can.

Statement of the Disappearance Form of the Identity Thesis

The central affirmation of DFIT is that sensations are nothing but physical processes (e.g., neurological and electro-chemical brain events). Rorty wishes to affirm this without making linguistic category mistakes, as well as declaring that linguistic schemes and analysis may not dictate what empirical science can or cannot find out about "what there is." Remember also that DFIT is intended as a strong form of materialism which does not yield to mere correlation between sensations and neurological/brain processes. The statement of DFIT must reflect this distinctive emphasis.

Unfortunately Rorty does not provide his own summary exposition of the minimal and key points of his DFIT, even though such an exposition is prerequisite to testing the sensibility or plausibility of the position. There is reason to believe, though, that we can summarize Rorty's position for ourselves without sacrificing fairness or adequacy. The following elaboration incorporates the distinctive emphases above, and the basic tenet of DFIT is entailed by the premises offered. It is thus formally adequate and does Rorty no injustice. DFIT may be summarized in this fashion:

1. Linguistic conventions and convenience have no relevance for ontological determinations.
2. The meaning or sense of sensation terms [hereafter ST] is not identical with the meaning or sense of certain physicalistic expressions [hereafter PE] (i.e., the language of physics, using terms like 'mass,' 'momentum,' 'electrical charge,' 'electron,' 'shape,' 'speed,' 'spatial location,' etc.)
3. Between competing accounts of some phenomenon, the one with greatest simplicity (i.e., countenancing fewer entities) and least burdened with problems (i.e., unanswered questions) is the acceptable alternative, thereby falsifying the more complex and problematic account.

4. As natural science advances we can, in principle, eliminate the referring use of ST in favor of PE without diminishing anything (such as ability to report, describe, explain or predict what there is) except linguistic convenience.
5. (Thus) The referent of ST is strictly identical with the referent of certain PE.
6. (Thus) There are no sensations (i.e., there are no other determinations than those employed by natural scientists).
7. Therefore, the referent of ST is identical with physicalistic states or processes.

It is not hard to see that 2.7 follows from 2.5 and 2.6 by an implicit disjunctive syllogism. The referent of ST and PE is the same, as 2.5 indicates, and this referent is either a sensation or physicalistic state/process. But since there are no sensations, according to 2.6, then the referent of ST and PE would have to be physicalistic. So then, Rorty's position has not been stated in a way which immediately makes it vulnerable as a non-sequitur. DFIT follows from two of its premises.

The other premises are crucial to DFIT as well. 2.1 guarantees that linguistic philosophy is not granted a magisterial position to restrict empirical discovery (particularly the advance hoped for in 2.4). 2.2 guards Rorty's position against charges of category mistake. Rorty openly denies that all attributes meaningfully predicable of sensations are also meaningfully predicable of physicalistic states/processes. The identity which Rorty wishes to assert, then, is not connotative but denotative (cf. 2.5). 2.2 also relieves Rorty of any need to produce adequate translations of ST into PE; he admits that ST (or mind talk) is meaningful, but he does not equate it with PE (or brain talk) with respect to meaning.

Premise 2.3 is the guiding principle which allegedly justifies moving from 2.4 to 2.6. The nail upon which the entire position hangs is premise 2.4, claiming that the purposes which were formally served conveniently by the use of ST could inconveniently become served by PE. From 2.4 Rorty derives the corollaries 2.5 and 2.6 (which in turn entail the conclusion, 2.7). 2.5 avoids the pitfall of mere correlation, and 2.6 is crucial to the materialist viewpoint. We can take the above exposition, then, as an adequate statement of Rorty's DFIT. But as it stands it is unconvincing.

In DFIT the premise which is crucial to the materialist view, 2.6, cannot be established in a way which is consistent with DFIT.

In DFIT Rorty argues that the referent of ST is strictly identical with the referent of certain PE (2.5). As it stands, though, this premise could just as well support the refutation of materialism - as when we conjoin to it the premise that 'The referent of ST is identical with sensations' (rather than physical states/processes). In that case the referent of PE themselves would turn out to be sensations! Wouldn't that turn the tables on the materialist? When the physicist is expressing certain scientific findings, it would turn out that he is talking, not about physical states or processes in the external world, but about his own personal (or somebody else's) sensations. Since sensations are mental phenomena, the "physical world" - including the brain and nervous system of man - which is analyzed by the natural scientist would in actuality be mental in nature (not physical) - opening the door to panpsychic and idealistic conclusions about reality. This shows us what a dangerous premise 2.5 is. Instead of having psychology reduce to physics, we would have physics reduce to psychology! Rorty can salvage his materialism only if he adds to 2.5 a denial of sensations (2.6).

1. The crucial premise cannot be established by the truthful substitution of PE for ST because this would leave mere correlation between sensations and physicalistic states/processes.

How might Rorty establish the non-existence of sensations? He might demonstrate that ST can be eliminated in favor of PE because some physicalistic expression could always be substituted *salva veritate* for the assertion mentioning a sensation. And of course this situation could very well obtain if PE and ST refer extensionally to the same thing. However, substitutability *salva veritate* and extensional agreement would not demonstrate that there exist no sensations. (As per Quine: 'creature with a heart' may indeed be substitutable *salva veritate* with 'creature with a kidney' without implying that kidneys do not exist.) It might just be the case that PE and ST refer to the same thing, and that this thing contains both physicalistic and sensational aspects. If sensations are always correlated with physicalistic states/processes, we would have an explanation of the substitutability *salva veritate* without needing to deny sensations at all. But mere correlation is not adequate for the position of maximal materialism, which closes this path for Rorty.

2. The crucial premise cannot be established by the functional substitution of PE for ST because the desired conclusion cannot be derived without violating the premise that linguistics must be ontologically neutral.

Rorty could now attempt to establish that there are no sensations (2.6) by pointing to a peculiar feature of our language and possible speech habits. He might, as premise 2.4 indicates, argue that all the linguistic functions of ST can be replaced by PE without sacrificing anything but convenience of expression. In that case it is theoretically possible that advances in natural science will lead us to abandon all ST (mind talk) and exclusively use PE (brain talk). The *faux pas* here is that such a consideration, used to establish the metaphysical conclusion that there are no sensations, violates Rorty's own insistence (2.1) that linguistic conventions and convenience have no relevance for ontological determinations. The line of thought under consideration moves from what is true about our language to what is true about what there is. If that tactic is denied to Rorty's opponents (who relish refuting maximal materialism by appeals to its resultant category mistakes given present linguistic usage), it must in fairness be denied to Rorty as well (who appeals to future

linguistic usage). The attempt to ground 2.6 in 2.4 is precluded by 2.1. Linguistic habit must remain ontologically irrelevant.

3. The crucial premise of DFIT is mediated by another premise which has no ontological relevance.

As indicated above, premise 2.6 in DFIT is derived from 2.4 in conjunction with the criterion stipulated in premise 2.3 (traditionally denominated "Occam's razor"). The soundness of 2.6 thus depends upon the truth or authority of those premises upon which it rests. On the hypothesis that all ST could be eliminated by PE, does the simplicity of PE really warrant the metaphysical conclusion that in reality there are no sensations? Notice here that "Occam's razor" is not a statement of fact, but a directive - an exceedingly generalized directive at that. It assumes the quite questionable cosmological law that nature is always characterized by economy and simplicity, never superfluity or unnecessary complexity. (Does anyone have enough empirical experience to generalize in such a grandiose fashion?) There seems to be plenty of reason to believe that the natural realm is not at all simple, a realm where everything is done in the most efficient and least complex, least mysterious fashion conceivable (e.g., methods of conception, as with salmon). One can wonder, then, whether the principle of non-superfluity is faithful to the actual state of affairs in the world. If it is not, then it cannot reliably be utilized to determine what the metaphysical situation actually is. Rorty can establish premise 2.6 only by resting upon a principle which is ontologically irrelevant. Simplicity and economy have nothing to do with whether things (like sensations) are real or not.

4. The crucial premise, 2.6, is unwarranted or gratuitous, being grounded in a claim which is false - the claim that sensation terms are unnecessary for correct reporting and describing of what there is.

Rorty maintains that the elimination of ST from our language would have no effect upon our ability to describe things properly. With the elimination of ST we might as well drop any thought of sensations themselves. It might be inconvenient to substitute PE for ST, but the fact that the substitution can be accomplished shows how unnecessary it is to suppose that sensations exist. Here is where we find the substantial backing for Rorty's DFIT. It only remains now for natural science to carry out the task of eliminating ST in favor of PE. That done, we will conclude that there are no sensations. What Rorty alleges, you see, is that as natural science advances there will no longer be any need to talk about mental states (e.g., hearing the "Star Spangled Banner," perceiving the orange hues of sunset, the taste of soy sauce, or knowing what day this is), for we will do just fine talking about the determinations of physics. Reports about mental states (such as perceptions) will become superfluous and could be replaced by purely physical analysis (electron and neuron talk).

1. The descriptive shortcoming of physicalistic expressions.

Rorty is just mistaken, I believe, in imagining that we lose only convenience if ST were dropped from our language. ST are indispensable to describing the observable qualities of things. Without observation terms we would not be able to describe the object of our experience fully; for instance, we would leave out the clearness and wetness of water, for no theoretical sentence using PE can analyze that descriptive truth.

PE could replace the ST used in describing features about things which are directly observed only if the PE were synonymous with the ST. PE might non-synonymously *explain* and/or *predict* phenomena reported by ST, but PE could *describe* the same phenomena as ST only if PE were synonymous with ST. Without synonymy, the two sentences would be used to assert different things about the phenomena in question. The two sentences might provide a way to speak of different aspects of the same phenomenon, but if those sentences have different senses, they would not be describing the same aspect of that thing, namely its observable quality.

A physicalistic statement does not indicate the observable appearance of things, even if it is capable of explaining how the appearance is mediated, stimulated, or recorded in the brain. No physicalistic expression renders the same description of a phenomenon which the sentence with sensation terms does *because* - as Rorty admits in premise 2.2 - PE and ST do not have the same connotation or sense. Accordingly a brain process and a mind process do not have all attributes in common (e.g., the intentionality or non-spatiality of the mental process). This dissimilarity indicates that PE cannot replace *all* the functions of ST, and in particular PE cannot offer a complete description of the appearance of things.

If ST were to be dropped, we would no longer be able to describe our sensory experience. What we are aware of sensing (e.g. the redness or tartness of the apple) is not described by PE, for PE describes a state or process with neurophysiological, chemical, electrical features and the like - features which are completely different from those which we report about our sensations.

Indeed, people can accurately describe their pain (e.g. as burning or throbbing) without knowing the first thing about neurophysiology or brain processes. So while the neurologist and his patient may be talking about the same thing (*viz.* the pain experience), they do not describe it in the same manner. The most complex and sophisticated scientific analysis of a phenomenon will never be equivalent to "ouch!" If the physician were to suggest to the patient who has just exclaimed his pain that "what he means" is some scientific description of C-fibers, neurons, synapses, etc., the patient would readily decline the translation. What he was talking about may be relevant to neurons, etc. in some fashion, but he was quite clearly not *talking about them*, but about his pain.

As long as we take premise 2.2 in DFIT seriously, what is spoken of in premise 2.4 can never take place. Sensation terms do not describe brain processes (not for most of us most of the time, anyway). Thus physicalistic terms which describe states and processes in the brain cannot replace the sensation terms. PE could replace ST only if they were synonymous, and they are not.

2. The implausibility of DFIT in saying what our describing is of

In order to escape the force of the above criticism Rorty must attempt to harmonize his theory with our practice of describing things in our experience (e.g. burning pains, zebras, childhood memories, musical tunes, etc.). He argues that when we use ST to directly report a sensory experience (like a pain) we are really giving a *non-inferential report about something in our brain or nervous system* (like a C-fiber firing). Mind you, we are not aware of any of the physicalistic determinations spoken of by the neurologist, but they are still what we directly report, according to Rorty. (As Thomas Hobbes held of old, what passes for a description of the external physical world of bodies is but the projection of one's internal physiology.)

The obvious defect with this claim is that I know how to give a direct report of physicalistic processes in the brain - that is, I know what procedures would have to be followed in order to put me in a position to observe the processes of the brain, and I know what kinds of terms I would use, as well as what kind of visual impressions I would associate with those terms, etc. However, when I see a zebra or taste a dill pickle or experience a throbbing pain, I use sensation terms (instead of physicalistic expressions) to describe them. Now, I realize the difference between the ST and PE, and I willingly, self-consciously choose the ST instead of the PE. When having the experience I am introspectively aware of my perceptions and their sensation determinations (i.e., the observable qualities of things) but of no physicist's or neurophysiologist's determinations whatsoever. The proponent of DFIT is at a loss to explain why I ordinarily use ST and only rarely use PE, those rare occasions being when I observe brain surgery, discuss books about subatomic particles, or like matters. It is beyond credibility that I - that nearly all linguistic users - should be so pervasively wrong about the things we think we are describing. When I describe the *taste* of a dill pickle, neither my intension nor my intention have anything to do with my brain. Rorty's attempt to say what my describing is of surely seems to be an *ad hoc* attempt to rescue an absurd thesis.

Conclusion

As is often the case, we are better to follow our common "sense" than the philosopher's speculations.

The disappearance form of the mind-body identity thesis is not initially plausible. It seems that Rorty was not warranted in advising that we must wait for the verdict of the scientific establishment before deciding whether the materialist view is correct or not. The identity thesis is afflicted with the kind of faults which empirical research cannot relieve. Because DFIT may be discounted in advance of further scientific developments, the worldview of materialism should not be granted a philosophical stay of life.

[1] Dr. Bahnsen holds the Th.M. from Westminster Theological Seminary and the Ph.D. in philosophy from the University of Southern California; he is presently the Resident Scholar at the Southern California Center for Christian Studies. This paper was originally presented at the Evangelical Philosophical Society, meeting in Philadelphia on December 28, 1976; it is presented here with minor changes.

[2] *The Review of Metaphysics* XIX (1965): 24-54