

THEORETICAL IDENTITIES AS EXPLANANTIA AND EXPLANANDA

Kevin Morris

I. INTRODUCTION

The mind-brain identity theory, the thesis that sensations are identical with properties or processes of the brain, was introduced into contemporary discussion by U.T. Place, Herbert Feigl, and J. J. C Smart in the 1950s.¹ Despite its widespread rejection in the following decades, the identity theory has received several carefully articulated defenses in recent years.² Aside from developing novel responses to well-known arguments against the identity theory, contemporary identity theorists have argued that the epistemological resources available to support the adoption of identities are more plentiful than has often been supposed; further, they have argued that mind-brain identities allow for the resolution of otherwise intractable explanatory puzzles about the phenomenal properties of experience. From an epistemological perspective, identity theorists have argued that a central reason for believing theoretical identities—both mind-brain identities as well as more mundane identities, such as $\text{water} = \text{H}_2\text{O}$ —stems from the explanatory power of the identities.³ On the other hand, often in response to “explanatory gap” worries about the relationship between the phenomenal and the physical, identity theorists have maintained that identities cannot themselves be explained; indeed, they have maintained that

it is hardly intelligible to request an explanation for an identity.⁴

The aim in what follows is to consider the role of theoretical identities as explanantia and explananda suggested in recent discussions of the identity theory and, in particular, to draw out a tension between these two roles. The central claim will be that insofar as they insist on the explanatorily efficacy of identities, identity theorists should concede that there is a good sense in which theoretical identities can be explained. On the other hand, however, it will be argued that this concession may not be especially problematic for the identity theorist.

II. IDENTITIES AS EXPLANANTIA

A pressing question for the identity theorist has always been why mind-brain identities should be endorsed. While contemporary identity theorists have generally rejected David Lewis’s contention that mind-brain identities follow as a matter of conceptual necessity from scientific discoveries about the causal roles of physical and neurophysiological properties,⁵ they have maintained that the prospects for justifying mind-brain identities by appealing to theoretical considerations are better than has sometimes been supposed. In particular, they have argued that not only do identities provide a simpler ontology than dualistic alternatives, but that identities can

also play a significant role as explanatory premises. Perhaps the clearest expression is provided by Ned Block:

The reason to think that the identities are true is that assuming them gives us explanations that we would not otherwise have and does not deprive us of explanations that we already have or raise explanatory puzzles that would not otherwise arise.⁶

In other words, mind-brain identities, and theoretical identities more generally, should be endorsed, at least in part, on the grounds that they provide better explanations than would otherwise be available. Assuming that inference to the best explanation is a viable epistemic principle, this provides a positive reason to endorse identities.

What do these explanations look like? Three possibilities can be discerned. First, it has been suggested that mind-brain identities explain correlations between the phenomenal and the physical—that the identity $\text{pain} = \text{C-fiber stimulation}$ explains the correlation of pain and C-fiber stimulation.⁷ Second, it has been claimed that identities allow for the best explanation of facts about causal-nomological relations between putatively higher-level properties and phenomena. Block, for instance, suggests that the identities $\text{temperature} = \text{mean molecular kinetic energy}$, $\text{water} = \text{H}_2\text{O}$, and $\text{freezing} = \text{lattice formation}$ should be endorsed on the grounds that together they provide the best explanation for why decreasing the temperature of water causes freezing, given an explanation for why decreasing the mean molecular kinetic energy of H_2O causes lattice formation.⁸ This can be represented schematically as follows:

II

T

R ($P_1 \dots P \dots P_n$)

$P_1 = M_1$, $P = M$, and so on.

R ($M_1 \dots M \dots M_n$)⁹

The idea here is to explain a causal or nomological relation between putatively higher-level properties and phenomena by first explaining that relation between lower-level properties and phenomena from a lower-level theory T_p and then identifying each putatively higher-level property or phenomenon with the appropriate lower-level property or phenomenon.

Third, and finally, it might be claimed that identities provide explanations of truths about the instantiation of higher-level properties or the occurrence of certain phenomena. David-Hillel Ruben provides the following example:

I can explain a gas's having a certain temperature t by its constituent molecules having mean kinetic energy m , and I can explain a change in a gas's temperature by a change in the mean kinetic energy of its constituent molecules. We explain in these cases, not just by laws of the coexistence of two types of phenomena, but by property or type-type identities.¹⁰

The claim, then, is that insofar as identities function as explanatory premises in these contexts, this provides a reason for thinking that they are true. Before considering the corresponding contention that identities cannot themselves be explained, it will be helpful to highlight an important thesis regarding the conditions under which identities can serve as explanatory premises. In particular, it is plausible that identities can only explain if the relevant explanatory contexts are construed "opaquely" in the sense that they are taken to relate representations of facts, rather than facts themselves. The reason is that insofar as identities are true, they cannot secure a transition from one fact to another distinct fact. And it is precisely this approach that Block takes in response to Jaegwon Kim's claim that rather than explain, identities merely allow the "re-description" lower-level facts using a different vocabulary.¹¹ Block writes:

Explanation is usually thought of as determining an "opaque" context. Just as knowledge of

the fact that freezing happened is not knowledge of the fact that lattice-formation happened, so also an explanation of the fact that freezing happened is not an explanation of the fact that lattice-formation happened.¹²

He continues:

In the transparent sense of “explain,” Kim is right and in the opaque sense he is not. And that is enough for my point: in one sense of “explanation,” the identities allow explanations one would not have without them.¹³

In other words, Block takes Kim to reason that if the identities are true, and if there is already an explanation of the fact of interest (say, an explanation for why decreasing the mean molecular kinetic energy of H₂O causes lattice formation), there is no further fact for identities to explain. But if explanations are taken to be sensitive to how facts are represented, having an explanation for a fact under one representation does not imply having an explanation of that fact under a different representation or conceptualization (say, an explanation for why decreasing the temperature of water causes freezing).

This does not conclusively establish that identities are explanatory in the contexts of interest, since this would require specifying sufficient conditions for explanatory success. It also does not establish that identities provide the best explanation in the relevant contexts, since this would require specifying the alternative explanations available and the conditions of explanatory goodness under which the identity explanation is superior. But insofar as the relevant explanatory contexts are taken to be sensitive to how facts are represented, there is a coherent account of how identities may possibly be capable of serving as explanatory premises, and so may potentially be justified by an application of inference to the best explanation.

Below a related thesis will be advanced about how identities themselves have to be understood if they are to play an explanatory

role. But this will be in the context of the claim that identities cannot themselves be explained, and it is this claim, and its role as a response to the “explanatory gap” problem, that will now be examined.

III. IDENTITIES AS EXPLANANDA

The suggestion that theoretical identities cannot be explained first explicitly appears in Kim’s “On the Psycho-Physical Identity Theory.” Here Kim writes:

A correlation statement cries out for an explanation: Why is it that whenever and wherever there is water, there is H₂O? Why is it that whenever and only whenever a person has pain he is in some specific brain state? Now, according to this line of reasoning, we can answer these questions if, and perhaps only if, we accept the corresponding identity statements. That is, we shall answer: Because water is H₂O, because pain is brain state B, and so on. But how can we explain these facts of identity? The answer is that they are not in need of explanation, that they cannot be explained. . . . It is nonsense to ask for an explanation of why Cicero is Tully, or why the Evening Star is the Morning Star; it is equally nonsensical to ask for an explanation of why water is H₂O, or why pain is brain state B. Water just is H₂O, and pain just is brain state B.¹⁴

Most philosophers have agreed with Kim on this point.¹⁵ And more recent defenses of the explanatory basicness of identities have often been advanced in response to the “explanatory gap” problem about the relationship between the phenomenal and the physical according to which there is a philosophically significant sense in which this relationship is not, and perhaps cannot, be adequately grasped or understood.¹⁶ The identity theorist’s response goes as follows. It is first noted that to say that there is an explanatory gap between the phenomenal and the physical seems to presuppose that the physical is distinct from the phenomenal. But if the identity theory is true, questions like, “How does the physical give rise to the

phenomenal?" have a false presupposition. And if it is maintained that identities cannot be explained—indeed, that it does not even make good sense to request an explanation for an identity—there would seem to not be any logical space left for an explanatory gap to inhabit.

Now, David Chalmers and Frank Jackson have responded to this strategy by maintaining that identities generally can be explained, and thus that phenomenal-physical identities fail to close the explanatory gap.¹⁷ The problem is that Chalmers and Jackson presume that identities can be explained in the sense that they follow from lower-level theory together with a priori functional analyses: for instance, the identification of water with H₂O can be explained in the sense that it follows from the empirical claim that H₂O is the clear, odorless liquid in the environment and the a priori analyses of water as the clear, odorless liquid in the environment.¹⁸ But contemporary identity theorists typically deny that such a priori analyses are available and, regarding the epistemology of theoretical identification, they have appealed to the explanatory efficacy of identities precisely because they think that theoretical identities like water = H₂O cannot be deduced, a priori, in the manner suggested by Chalmers and Jackson.¹⁹ And thus, identity theorists will hardly be persuaded by Chalmers and Jackson on this point.

Yet this only establishes that theoretical identities cannot be explained in a certain

manner. Similarly, while Robert Causey notes that identities cannot be given causal or mechanistic explanations,²⁰ it only follows from this that identities cannot be explained if it is presumed that all explanation is causal or mechanistic. And upon reflection, it is doubtful that the identity theorist should insist on the explanatory basicness of identities as the thesis that there is no sense in which any identity can be explained.

An intuitive case for this goes as follows. The identity theorist takes instances of the following schema to count as explanations and contends that the identity may serve as an explanatory premise in the explanation:

I2

S has P at t.
P = M
Thus, S has M at t.

For instance, that an individual is in pain can be explained by noting that the individual has C-fiber stimulation and that C-fiber stimulation = pain. But in this case, it is hard to see why instances of the following schema cannot count as explanations:

I3

M = P
P = Q
Thus, M = Q

Why should the identity in instances of I2 be taken to explain, but not the identities in instances of I3? For instance, it is not clear why any of the following may not count as explanations:

<p>I3' Hesperus = Venus Venus = Phosphorus Thus, Hesperus = Phosphorus</p>	<p>I3'' Mark Twain = Samuel Langhorne Samuel Langhorne = Sam Clemens Thus, Mark Twain = Sam Clemens</p>	<p>I3''' Water = that stuff That stuff = H₂O Thus, water = H₂O</p>
---	--	---

It was noted above that insofar as identities can explain, this explanatory function consists in allowing a transition from one conceptualization of a fact to another, even if they are different conceptualizations of the same fact. But if the identity can figure as an explanatory premise in instances of I2 in virtue of this role (as well as in the other identity explanations sketched above, including I1), the identities in I3 would seem to play this same role in virtue of getting a new identity from two antecedent identities. It is no help to contend that there is no transition in I3 from one fact to a new, distinct fact—perhaps it will be maintained that all identities express the trivial fact that something is what it is—since insofar as the identity in an instance of I2 is true, this will be the case here as well.

It is worth considering the idea that theoretical identities like $\text{water} = \text{H}_2\text{O}$ simply express the fact that something is what it is in a bit more detail, since this may provide the most substantial reason that identity theorists have offered in favor of the explanatory basicness of identities. David Papineau, for instance, maintains that if a conscious property is a material property, “then there is no mystery of why it is what it is” and that “you cannot explain why one thing is itself.”²¹ Likewise, Michael Tye writes that if an identity is true, then “only one state exists, conceived of in two ways, and that state must be self-identical,”²² while Joseph Levine insists that “things are what they are; there is no sense in explaining that.”²³ The idea seems to be that a theoretical identity expresses the fact that a property or phenomenon is self-identical, that it is what it is. Since there is no explanation for why something is what it is, there is no explanation for why $\text{water} = \text{H}_2\text{O}$ or why $\text{pain} = \text{C-fiber stimulation}$.

Perhaps it is true that why something is what it is cannot be explained and that there is a sense in which identities express such trivial facts. But this construal of theoretical

identities cannot be what identity theorists have had in mind when taking identities to be explanatorily efficacious, at least because it is hard to see how the fact that something is what it is can explain anything at all in the relevant contexts: insofar an explanation for why decreasing the temperature of water causes freezing can consist, in part, in the identity $\text{water} = \text{H}_2\text{O}$, that $\text{water} = \text{H}_2\text{O}$ expresses the fact that water is what it is would seem to play no role whatsoever in the success of this explanation. Indeed, by focusing on the suggestion that theoretical identities express the fact that something is what it is, it could just as well be argued that since the fact that something is what it is epistemically trivial and explanatorily idle, it follows that theoretical identities are epistemically trivial and explanatorily idle. The identity theorist does not want this conclusion—it raises the question of why a substantive epistemology for theoretical identities is needed at all and, moreover, how the explanatory efficacy of identities can play this role. More generally, the supposition is that theoretical identities can play a substantive explanatory role, a role substantive enough to warrant the adoption of epistemically nontrivial identities.

This goes some distance toward showing there has to be a sense in which the identity theorist supposes that theoretical identities do more than simply express trivial facts. But when identities are understood in terms of their role as epistemically nontrivial explanatory premises, and the relevant explanatory contexts are taken to be sensitive to how facts are represented, it is not clear why an identity cannot be explained by appealing to other identities. And if identities can be explained, the identity theorist should not maintain that it does not make sense to request an explanation for an identity.²⁴

Are there any other arguments in favor of the explanatory basicness of identities? It is sometimes argued that the necessity of

theoretical identities precludes their explainability. That is, following Saul Kripke,²⁵ contemporary identity theorists have maintained that identities are necessary if true; and they have further argued that this precludes the explainability of theoretical identities. Thus, Papineau contends in this context that a single person could not possibly have been different people,²⁶ while Andrew Melnyk writes that “it seems misconceived to ask why $a = b$ (when it is)” and that this is so “because identities hold necessarily, and it makes sense to explain why something is the case only if it might not have been the case.”²⁷ On this approach, necessary truths generally cannot be explained, and the explanatory basicness of theoretical identities is an instance of the explanatory basicness of necessary truths.

This line of thought is not very promising, for while necessary truths may not admit of certain forms of explanation—for instance, causal explanations—it is doubtful that the general explainability of necessary truths should be denied. It is usually recognized that mathematical truths can be explained, and insofar as mathematical truths are necessary, a sense in which necessary truths can be explained should be countenanced.²⁸ And within the metaphysics of mind, it is presumed to be incumbent upon the physicalist to explain why, if at all, instances of mental properties are metaphysically necessitated by how things are physically—that is, why there are necessary connections between how things are physically and the distribution of mental properties in the world.²⁹ Further, insofar as the identity theorist takes metaphysically necessary mind-brain identities to explain mind-brain correlations, the identity theorist would seem to be committed to the explainability of necessary truths: as Kim notes, if it is necessary that pain is C-fiber stimulation, then the correlation of pain with C-fiber stimulation should likewise be regarded as necessary.³⁰ But since the identity is taken

to explain the correlation, it follows that the identity theorist is committed to necessary truths being explainable. Finally, $I3'$ – $I3''$ each involve paradigm cases of identities that are necessary if true (both as explanantia and as explananda). But this does not seem to preclude them from counting as explanations.³¹

It might be claimed that even if the explanatory basicness of identities cannot be maintained for the reasons just sketched, there is nonetheless something suspicious about instances of $I3$.³² For one, it might seem odd to say that water is H_2O because water is that stuff and that stuff is H_2O , or that Mark Twain is Sam Clemens because Mark Twain is Samuel Langhorne and Samuel Langhorne is Sam Clemens. But the force of this line of thought is questionable. It is plausible that any oddness regarding explanations of identities that take the form of $I3$ will arise with any explanation that appeals to an identity as a premise. For instance, it may seem odd to say that decreasing the temperature of water causes freezing because decreasing the mean molecular kinetic energy of H_2O causes lattice formation and $\text{water} = H_2O$, $\text{temperature} = \text{mean molecular kinetic energy}$, and $\text{freezing} = \text{lattice formation}$ (an instance of $I1$); and it may seem odd to say that water is present because H_2O is present and $\text{water} = H_2O$ (an instance of $I2$). But the identity theorist supposes that these are explanations nonetheless, and any oddness present here would seem to be the same oddness associated with the because-statements generated from instances of $I3$. In this way, the present worry does not provide a reason for thinking, as identity theorists have, that identities can explain but cannot be explained.

It may also be noted that the contention that instances of $I3$ are genuine explanations has the following consequence. Schematically, if $M = Q$ can be explained by noting that $M = P$ and that $P = Q$, it would also seem that $M = P$ can be explained by noting that

$M = Q$ and that $P = Q$, and that $P = Q$ can be explained by noting that $M = P$ and that $M = Q$. And it might seem that this is intolerable, perhaps because it shows that putative explanations of identities fail to exhibit a certain sort of asymmetry: if an identity $M = Q$ can be explained by the identities $M = P$ and $P = Q$, that same identity $M = Q$ might also serve as a premise in an explanation for why $P = Q$ or why $M = P$.

It is plausible that this is indeed a consequence of the claim that instances of I3 are explanations. Further, it should be conceded that this consequence may call into question the depth of explanations of identities that appeal to other identities as premises (an issue that will arise in section IV). But it is doubtful that this observation can be utilized to support the contention that identities can explain but cannot be explained, since it is doubtful that it can call into question the explanatory credentials of instances of I3 without generally calling into question the explanatory credentials of explanations that appeal to identities as premises. This is because insofar as the present line of thought threatens the explanatory import of instances of I3, this would seem to turn on the observation that the conclusion of such an explanation may, in another context, serve as a premise in an explanation of one of *its* explanatory premises. The problem is that the same worry can be raised quite generally against explanations that utilize identities as premises. If the presence of water can be explained by noting that H_2O is present and that $\text{water} = H_2O$, the presence of H_2O can likewise be explained by noting that water is present and that $\text{water} = H_2O$. But since the identity theorist maintains that the presence of water can in fact nonetheless be explained by noting that H_2O is present and that $\text{water} = H_2O$, the identity theorist should not reject the contention that instances of I3 may count as explanations for the reason here under consideration.

Finally, it may be claimed that while certain identifications may be explainable and so may admit of explanatory demands, there are principled reasons for thinking that phenomenal-physical identities do not. In this way, the identity theorist could continue to endorse the basic contention that if there is reason to believe that phenomenal-physical identities are true, there is no space for an explanatory gap. In particular, it might be claimed that phenomenal properties are represented under phenomenal concepts, which are not a priori associated with any descriptive expressions, and that this somehow precludes the explanation of mental-physical identities.³³

Yet aside from worries about the claim that concepts of the phenomenal can indeed be so understood, as well as worries about the extent to which there can be such a unique kind of conceptualization for something that is physical,³⁴ the appeal to phenomenal concepts at best establishes that phenomenal-physical identifications cannot be explained by appealing to a priori analyses. Moreover, if it is presumed that phenomenal-physical identifications are epistemically nontrivial and can potentially be justified by appealing to their explanatory efficacy, according to the line of thought advanced above it follows that there is a sense in which requests for an explanation for such identifications are legitimate. More generally, the identity theorist should not argue for the explanatory basicness of identities by noting the lack of available a priori analyses, since a priori analyzability does not seem to determine the explainability of an identity.

How closely should the explanatory efficacy of theoretical identities be bound to the explainability of such identities?³⁵ An especially strong view would be that the explanatory efficacy of an identity entails that the identity can be explained. But the considerations just advanced may more modestly be taken to show that there is at least one sense

in which some theoretical identities can be explained. This falsifies the claim that there is no sense at all in which any such identity can be explained. These considerations suggest that insofar as an identity statement is explanatorily efficacious, it follows that it at least makes sense, or is intelligible, to request an explanation for that identity. There is an analogy here with laws of nature: generally, it makes sense to ask why a law that figures in successful explanations obtains. In many cases, the law will indeed be explainable; in other cases, an explanation may not be available presently or in the foreseeable future. But it will still make sense, or be intelligible, to ask why the law obtains.

IV. MUST THE IDENTITY THEORIST INSIST ON THE EXPLANATORY BASICNESS OF IDENTITIES?

If identities can be explained, an explanation that appeals to phenomenal-physical identities will not allow all potential explanatory demands regarding the phenomenal vis-à-vis the physical to be set aside. And in this way, the identity theorist should not claim that the identity theory solves the explanatory gap problem in virtue of the explanatory basicness of theoretical identities.

But is there any good reason to think that the viability of the identity theory turns on the contention that mind-brain identities cannot be explained? There are several reasons for doubting that this is the case. The first thing to notice is that it is not a general criterion for explanatory success that the explanatory premises should not themselves admit of legitimate explanatory demands. And in this way, that identities may themselves admit of explanatory demands in no way threatens the explanatory credentials of identities in the contexts noted above. Perhaps more significantly, theoretical identities can be seen to play a role in explanations of a quasi-reductive sort, and in this sense can be said to close

“explanatory gaps” even if they do not allow all subsequent explanatory demands to be set aside. In particular, insofar as mind-brain identities are true, mental truths will be explainable in physical terms without appealing to any laws relating mental properties to each other or relating physical properties to mental properties.³⁶ Consider again Block’s explanation of why decreasing the temperature of water causes freezing. In this explanation, a truth about water is explained by appealing to lower-level physical theory together with the requisite identities. While this may not count as a reductive explanation under all accounts of what reductive explanation requires—according to one such conception, reductive explanation of the phenomenal is possible just in case there are a priori, broadly definitional connections between the physical and the phenomenal³⁷—it plausibly provides for a sense in which a truth about water can be reductively explained. And in this way, mind-brain identities can close explanatory gaps between the physical and the phenomenal even if the identities admit of explanatory demands.

The explanatory nonbasicness of identities should only worry the mind-brain identity theorist if this explanatory nonbasicness provides a reason for thinking that mind-brain identities are false. But it is not clear how such an argument would go. Indeed, if it is maintained that identities generally cannot be explained, the apparent intelligibility of explanatory demands aimed at phenomenal-physical identities becomes quite problematic for the identity theorist. Thus, for instance, Levine maintains that whenever an identity might seem to admit of explanatory demand, the target explanandum should be construed in such a way that it turns out to be something other than an identity. Thus, he writes:

Of course there are identity claims that one can seek explanations for, but they always turn out to be, if not requests for evidence, ques-

tions about how or why distinct properties are coinstantiated. So, for instance, I can express wonder that this full-grown man I am now facing is the same person as the little boy I met 20 years ago, or even that this apparently continuous liquid I call “water” could be the same thing as a collection of H_2O molecules. But in both cases it’s clear that what I’m wondering about is how the very same object could instantiate these very different properties.³⁸

Levine concedes that explanatory demands which target phenomenal-physical identities are in fact intelligible (which he thus labels “gappy identities”); but given his commitment to the explanatory basicness of identities, this seems to imply that such phenomenal-physical identities are false or that such requests should be interpreted as requests for an explanation for how one thing could instantiate distinct properties. In this way, while the “gappiness” of phenomenal-physical identifications may not entail that they are false, Levine takes this to present a clear challenge for the physicalist, since he presumes that the intelligibility of such requests is best accounted for by presuming that the phenomenal is distinct from the physical.

This line of thought, however, crucially supposes that explanatory demands of true identities do not make good sense, and if the above line of thought is on track, this is a presumption that should be rejected. And the challenge, then, is to show that the sense in which phenomenal-physical identifications admit of explanatory demands somehow goes beyond the sense in which more mundane identifications like $water = H_2O$ admit of such demands. In this way, the concession that identities can be explained can be seen to alleviate otherwise pressing worries about the adoption of phenomenal-physical identities.

Finally, as briefly suggested in section III, even if there is a sense in which identities can be explained, and so admit of legitimate explanatory demands, the depth of these explanations may be questioned.

For according to the line of thought here defended, the sense in which mind-brain identities can be explained is just the sense in which an identity such as Mark Twain is Sam Clemens can be explained by noting that Mark Twain is Samuel Langhorne and that Samuel Langhorne is Sam Clemens. If the considerations advanced above are on track, there is no good reason to deny that this is an explanation, given an antecedent commitment to the explanatory efficacy of identities in the contexts of interest for the identity theorist. Nonetheless, it does not seem to be especially deep explanation. Consider, then, the following passage from Papineau:

I say that once you really accept that pain, say, really is some material M, then you will see that this requires no more explanation than does Mark Twain = Samuel Clemens. Identities need no explaining.³⁹

The considerations advanced above provide no reason for doubting Papineau’s claim that mind-brain identities require no more explanation than the identity Mark Twain = Samuel Clemens. Indeed, given the benign consequences of the explanatory nonbasicness of identities, it may be conceded that there is a sense in which Papineau is correct that identities need no explaining—explanations of identities are not especially deep and the lack of an explanation for an identity does not seem to provide a reason for doubting the truth of the identity. But it should not be claimed, as Papineau and others have, that identities need no explaining just in that they cannot be explained.

V. CONCLUSION

The identity theorist assumes that theoretical identities, such as potential mind-brain identities, are epistemologically problematic. The identity theorist thus needs an epistemology, and the strategy in recent years has been to appeal to explanatory considerations. But the identity theorist also

wants theoretical identities, while explanatorily efficacious, to be explanatorily basic. Yet there is a tension in this combination of views: identities are explanatorily nonbasic in the sense in which they are explanatorily efficacious, and they are explanatorily idle, and perhaps epistemically trivial, in the

sense in which they are explanatorily basic. Whether this observation seriously threatens the identity theorist, however, is not clear, and some reasons have been provided for doubting that this is the case.

Tulane University

NOTES

A version of this essay was presented at the 16th Annual University of Waterloo Graduate Conference. Thanks to the audience members on that occasion for helpful questions and discussion. Thanks also to Jaegwon Kim and Max Pines for helpful comments and discussion on the issues addressed in this essay. Finally, thanks to an anonymous reviewer for helpful comments and criticism.

1. Feigl 1958, Place 1956, Smart 1959.
2. Block 2002 and forthcoming, Block and Stalnaker 1999, Hill 1991, Hill and McLaughlin 1999, Papineau 2004, Polger 2004. Contemporary versions of the identity theory have often been advanced in the more general context of defending the view known as “a posteriori physicalism,” according to which (roughly) physicalism is true so long as all properties can be identified, a posteriori, with physical properties or properties realized by physical properties (and only by physical properties).
3. Block 2002 and forthcoming, Block and Stalnaker 1999, Hill 1991, Hill and McLaughlin 1999, and Melnyk 2003. Similar claims are advanced in Bechtel and McCauley 2001. While the focus here will be on physical identifications—the identification of, say, a mental property with a physical property—it has been suggested that essentially these same considerations can warrant the adoption of functional identifications, the identification of a mental property with a functional property (see Melnyk 2003). It is plausible that subsequent discussion applies to this latter suggestion as well, though the discussion will be framed in terms of physical identifications.
4. Block 2002, Block and Stalnaker 1999, Papineau 1998 and 2004, and Tye 2002.
5. Lewis 1966. The relevance of this for the claim that identities cannot be explained will be briefly considered in section III.
6. Block forthcoming.
7. Hill 1991, Hill and McLaughlin 1999.
8. Block 2002, Block and Stalnaker 1999. Melnyk 2003 endorses a similar strategy.
9. This schematization is adopted, with some minor modifications, from Kim 2005, pp. 143–144.
10. Ruben 1990, p. 219.
11. Kim 2005, chap. 5.
12. Block forthcoming.
13. Ibid. A similar line of thought regarding the opacity of explanatory contexts and the explanatory efficacy of identities is advanced in Ruben 1990, p. 219.
14. Kim 1966, p. 229.
15. Several years later, for instance, Robert Causey contended that identity statements can be distinguished from statements expressing nomological correlations precisely on the basis that the latter

admit of explanatory demands but the former do not (Causey 1977, chap. 2). Achinstein 1983 presents a number of alleged counterexamples to Causey's claim about the explanatory basicness of identities. However, Achinstein's examples all involve identities with definite descriptions flanking the identity sign, and contemporary mind-brain identity theorists have typically understood theoretical identities to involve rigid designators and thus would contend that Achinstein's examples do not involve genuine identities.

16. See Levine 1983. Similar claims are advanced in Nagel 1974 and McGinn 1989.

17. Chalmers and Jackson 2001, p. 354.

18. This manner of understanding theoretical identities is also defended in Jackson, Pargetter, and Prior 1982 and Lewis 1966.

19. See, for instance, Block and Stalnaker 1999.

20. Causey 1977, chap. 2.

21. Papineau 1998, p. 379.

22. Tye 2002, p. 33.

23. Levine 2001, p. 65.

24. It will not be much help for the identity theorist to appeal to the idea that the identity $\text{water} = \text{H}_2\text{O}$ means, or expresses the same proposition as, the identity $\text{water} = \text{water}$ on the grounds that the meaning of "water" just is its referent (H_2O). For in this case, all of the same questions raised about trivial identity facts can be reapplied to the distinction between the apparently trivial proposition expressed by $\text{water} = \text{H}_2\text{O}$ and the putative explanatory efficacy of $\text{water} = \text{H}_2\text{O}$. That is, just as the explanatory efficacy, and epistemic nontriviality, of $\text{water} = \text{H}_2\text{O}$ cannot easily be accounted for by the trivial fact that something is what it is, it would also seem that the explanatory efficacy of $\text{water} = \text{H}_2\text{O}$ is not easily accounted for by the suggestion that $\text{water} = \text{H}_2\text{O}$ expresses the same trivial proposition as the identity $\text{water} = \text{water}$ (or $\text{H}_2\text{O} = \text{H}_2\text{O}$): insofar as the identity $\text{water} = \text{H}_2\text{O}$ is explanatorily efficacy, it is not in virtue of expressing this proposition (if it does).

25. See Kripke 1971.

26. Papineau 1998, p. 379.

27. Melnyk 2003, p. 52.

28. See, for instance, Steiner 1978a and 1978b. While Resnik and Kushner 1987 offer a critical analysis of Steiner's account, they concede that there are senses in which mathematical truths can be explained.

29. See, for instance, Horgan 1993 and Melnyk 2003.

30. Kim 2005, pp. 135–139. It may be noted that Kim rejects the identity theorist's claim that identities can indeed explain correlations.

31. In addition to the arguments that identity theorists have explicitly advanced (the appeal to the triviality of identity facts and the necessity of identity), it might seem that the following line of thought (suggested by an anonymous reviewer) supports the explanatory basicness of identities. It is sometimes maintained that entities identical with respect to qualitative properties can nonetheless be distinct entities (see Black 1952). But in this case, given distinct entities with identical qualitative properties, the qualitative properties of the entities cannot be invoked to explain why the entities are indeed distinct, and so will not be able to explain identities involving the entities (say, the claim that this entity is entity A, that this [other] entity is entity B, and so on). The force of this argument, however, is unclear. The sense in which identities can be explained suggested by I3 does not seem to turn on whether indiscernibility with respect to qualitative properties implies identity. Moreover, it was conceded that there is a

good sense in which identity facts are indeed brute and inexplicable. But if the line of thought advanced in the text is on track, it is compatible with this that there is another sense in which identities can be explained.

32. The worries in this and the subsequent paragraph were raised by an anonymous reviewer.
33. Phenomenal concepts are defended, for instance, in Carruthers 2004 and Papineau 2004.
34. See Levine 2001 and Tye 2009 for the first sort of worry and Chalmers 2006 for the second.
35. This issue, in effect, was raised by Sanford Goldberg when a version of this essay was presented at the 16th Annual University of Waterloo Graduate Conference in May 2009.
36. That this qualifies as a sort of reductive explanation is suggested in Kim 2005, p. 105n17.
37. This demanding view of reductive explanation is defended in Chalmers 1996 and Chalmers and Jackson 2001.
38. Levine 2001, p. 65.
39. Papineau 2004, p. 150.

REFERENCES

- Achinstein, Peter. 1983. *The Nature of Explanation* (New York: Oxford University Press).
- Bechtel, W., and R. McCauley. 2001. "Explanatory Pluralism and the Heuristic Identity Theory," *Theory and Psychology*, vol. 11, pp. 736–760.
- Black, Max. 1952. "The Identity of Indiscernibles," *Mind*, vol. 61, pp. 153–164.
- Block, Ned. 2002. "The Harder Problem of Consciousness," *Journal of Philosophy*, vol. 99, pp. 1–35.
- _____. Forthcoming. "Functional Reduction," in *Supervenience in Mind: A Festschrift for Jaegwon Kim*, ed. T. Horgan, D. Sosa, and M. Sabates (Cambridge, Mass.: The MIT Press), <http://www.nyu.edu/gsas/dept/philo/faculty/block/papers/Kimfestschrift.pdf>
- _____, and R. Stalnaker. 1999. "Conceptual Analysis, Dualism, and the Explanatory Gap," *Philosophical Review*, vol. 108, pp. 1–46.
- Carruthers, Peter. 2004. "Reductive Explanation and the 'Explanatory Gap,'" *Canadian Journal of Philosophy*, vol. 34, pp. 153–174.
- Causey, Robert. 1977. *Unity of Science* (Dordrecht, The Netherlands: D. Reidel).
- Chalmers, David. 1996. *The Conscious Mind* (Oxford: Oxford University Press).
- _____. 2006. "Phenomenal Concepts and the Explanatory Gap," in *Phenomenal Concepts and Phenomenal Knowledge: New Essays on Consciousness and Physicalism*, ed. T. Alter and S. Walter (New York: Oxford University Press).
- _____, and F. Jackson. 2001. "Conceptual Analysis and Reductive Explanation," *Philosophical Review*, vol. 110, pp. 315–360.
- Feigl, Herbert. 1958. "The 'Mental' and the 'Physical,'" in *Concepts, Theories, and the Mind-Body Problem*, ed. H. Feigl, M. Scriven, and G. Maxwell (Minneapolis: University of Minnesota Press).
- Hill, Christopher. 1991. *Sensations: A Defense of Type-Materialism* (Cambridge: Cambridge University Press).
- _____, and B. McLaughlin. 1999. "There Are Fewer Things Than Are Dreamt of in Chalmers' Philosophy," *Philosophy and Phenomenological Research*, vol. 59, pp. 445–454.
- Horgan, Terence. 1993. "From Supervenience to Superdupervenience: Meeting the Demands of a Material World," *Mind*, vol. 102, pp. 555–586.
- Jackson, F., R. Pargetter, and E. Prior. 1982. "Functionalism and Type-Type Identity Theories," *Philosophical Studies*, vol. 42, pp. 209–225.

- Kim, Jaegwon. 1966. "On the Psycho-Physical Identity Theory," *American Philosophical Quarterly*, vol. 3, pp. 227–235.
- _____. 2005. *Physicalism, or Something Near Enough* (Princeton, N.J.: Princeton University Press).
- Kripke, Saul. 1971. "Identity and Necessity," in *Identity and Individuation*, ed. M. Munitz (New York: New York University Press).
- Levine, Joseph. 1983. "Materialism and Qualia: The Explanatory Gap," *Pacific Philosophical Quarterly*, vol. 64, pp. 354–361.
- _____. 2001. *Purple Haze* (Oxford: Oxford University Press).
- Lewis, David. 1966. "An Argument for the Identity Theory," *Journal of Philosophy*, vol. 63, pp. 17–25.
- McGinn, Colin. 1989. "Can We Solve the Mind-Body Problem?" *Mind*, vol. 98, pp. 349–366.
- McLaughlin, Brian. 2001. "In Defense of New Wave Materialism: A Response to Horgan and Tienson," in *Physicalism and Its Discontents*, ed. B. Loewer (New York: Cambridge University Press).
- Melnyk, Andrew. 2003. *A Physicalist Manifesto* (Cambridge: Cambridge University Press).
- Nagel, Thomas. 1974. "What Is It Like to Be a Bat?" *Philosophical Review*, vol. 83, pp. 435–450.
- Papineau, David. 1998. "Mind the Gap," *Noûs*, vol. 32, pp. 373–388.
- _____. 2004. *Thinking about Consciousness* (Oxford: Oxford University Press).
- Place, U. T. 1956. "Is Consciousness a Brain Process?" *British Journal of Psychology*, vol. 47, pp. 44–50.
- Polger, Thomas. 2004. *Natural Minds* (Cambridge, Mass.: The MIT Press).
- Resnik, M., and D. Kushner. 1987. "Explanation, Independence, and Realism in Mathematics," *British Journal for the Philosophy of Science*, vol. 38, pp. 141–158.
- Ruben, David-Hillel. 1990. *Explaining Explanation* (New York: Routledge).
- Smart, J. J. C. 1959. "Sensations and Brain Processes," *Philosophical Review*, vol. 69, pp. 141–156.
- Steiner, Mark. 1978a. "Mathematics, Explanation, and Scientific Knowledge," *Noûs*, vol. 12, pp. 17–28.
- _____. 1978b. "Mathematical Explanation," *Philosophical Studies*, vol. 34, pp. 135–151.
- Tye, Michael. 2002. *Color, Consciousness, and Content* (Cambridge, Mass.: The MIT Press).
- _____. 2009. *Consciousness Revisited: Materialism without Phenomenal Concepts* (Cambridge, Mass.: The MIT Press).

2