

ANTI-INDIVIDUALISM, MATERIALISM, NATURALISM

Tomáš Hříbek

ABSTRACT: This paper starts from the familiar premise that psychological anti-individualism is incompatible with materialism. It attempts to state more clearly what this incompatibility consists in, and – rather than arguing in detail for any particular resolution – to inquire whether this incompatibility admits any resolution. However, the paper does offer a conditional argument concerning the possibility that the incompatibility is genuine and cannot be resolved. Provided that anti-individualism and materialism cannot be squared, and anti-individualism is correct, it follows that materialism has to be abandoned. If so, the situation is not as disastrous as it might at first seem. We need not, in consequence of our inability to construe a coherent metaphysics of mind, give up on intentional vocabulary any more than we must stop, in consequence of our inability to make sense of induction, anticipating the future.

KEYWORDS: Anti-individualism, materialism, naturalism, Tyler Burge.

1. Introduction

Ever since Burge's classic statement of the issue,¹ the relationship between anti-individualism – otherwise known as externalism – and materialism has seemed strained, if not downright contradictory. And yet the tensions between these two doctrines have been discussed considerably less than apparent tensions between anti-individualism and self-knowledge, and between anti-individualism and mental causation. To be sure, I cannot do justice to the complex issue of the relations between anti-individualism and materialism in the space of a short paper.² My aim is mostly just to chart a conceptual geography of this area; to identify the main issues and options. I will, however, make a conditional argument:

¹ See Burge (1979, 110 – 111) and, more recently, Burge (1993, 105).

² I tried to analyze this relationship in detail in my unpublished doctoral dissertation, *The Metaphysics of Anti-Individualism* (University of Minnesota, Minneapolis, 2006). The present paper incorporates parts of the introduction and the final chapter of the dissertation.

if anti-individualism and materialism are incompatible; if, further, anti-individualism is correct about the way we individuate propositional events; and if, in consequence, materialism has to be abandoned; then the resulting situation does not necessarily call for some drastic change in our practice.

The following two sections will clarify the theses of materialism and anti-individualism, in particular their respective epistemic and modal claims, and explain why the two theses are incompatible. The fourth section places the issue of an apparent incompatibility between anti-individualism and materialism alongside the two aforementioned incompatibility issues, and explains why the former has so far attracted much less attention by means of a summary of various attempts at reconciling anti-individualism with different forms of materialism. The final, fifth, section offers a more detailed argument alluded to above: it suggests that we might, with respect to intentional vocabulary, be in a situation similar to our relationship to induction. We cannot make sense of induction, but it does not mean that inductive reasoning is irrational; likewise, we cannot build a coherent materialist metaphysics of mind, but it does not mean that we should stop using the vocabulary of propositional attitudes. It might be as natural for us to use intentional idioms as it is to reason inductively.

2. Materialism

Philosophers of mind converge at least on one thing today—most assume that some form of materialism *must* be correct. Let us see first what the thesis of materialism amounts to, at least in general terms; we can worry about the modal claim later. It is clear that contemporary materialism is only remotely related to the eighteenth-century doctrine, according to which there exists nothing but material substance. Rather, materialism is understood today as a view about which science is basic, in the sense of picking out the ultimate determinants of everything else. It is generally agreed that physics is such a basic science—hence, the term “physicalism” is often used as a synonym for “materialism” in contemporary debates. In this usage, then, physicalism is a recent descendant of the old materialism. Other authors, however, relate the two doctrines in terms of their respective scope, rather than their historical origins. Thus, materialism is regarded by some as a weaker, less ambitious doctrine,

according to which physical objects can instantiate some nonphysical properties. Others go even as far as allowing the existence of some nonphysical, abstract objects, such as numbers, laws, or propositions, apart from the physical realm. By contrast, physicalism is regarded as a narrower, radical theory, according to which all objects, properties and relations that exist are physical. Finally, some philosophers, especially in the mid-twentieth century, distinguished between materialism as a robust doctrine, on the one hand, and physicalism as a linguistic doctrine, on the other. In this usage, materialism was indeed a theory about which objects or properties ultimately exist – viz., material objects as opposed to, say, sense data – whereas physicalism was a thesis about which language should be treated as basic – viz., the language of physics as opposed to, say, the language of psychology.³

Now, let us ignore materialism's remote origins; the ontological status of abstract entities; as well as the fine distinction between robust and linguistic doctrines. As I see it, the story of modern materialism begins in the 1950s as a by-product of the debates about reduction in philosophy of science. The idea of physics as the basic science led to the attempts at reducing the so-called special sciences (e.g., chemistry, biology, economics, and psychology) to physics. As this goal proved unattainable, modern materialism got underway as a project of articulating the exact relation between the domains of special sciences, on the one hand, and the domain of physics, on the other, that would fall short of reduction. Thus, the reductive materialism of the 1950s and the 60s, which hoped to ultimately derive the generalizations of special sciences, including psychology, from the generalizations of physics, was followed by nonreductive materialism that admits that no such derivation is possible, requiring only that all particulars recognized by special sciences are physical particulars. This contrast between the two varieties of materialism has often been expressed in terms of the type-token distinction: reductive materi-

³ For a discussion of the differences between the early modern and contemporary materialisms, see Kim (1998). Robinson is one of those who distinguish between materialism and physicalism in terms of scope: see Robinson (1993). Physicalism as a linguistic version of materialism was introduced by logical empiricists. See esp. Neurath (1931) and Carnap (1932). For an up-to-date comprehensive defense of physicalism, see Poland (1994). Among the things discussed by Poland are various options in defining the term "physical." For example, it is not a trivial matter whether the definition is tied to current physics or, rather a future complete form thereof. Here I ignore the issues that might be raised by this controversy.

alism is a type-identity theory, since it believes in the derivations of higher-level types from the physical types, while non-reductive materialism is a token-identity theory, as it requires no more than that all higher-level, in particular psychological, tokens are physical tokens.⁴ As for numbers, propositions and such, I believe their abstract nature is compatible with materialism, as long as these entities are considered to be outside of the natural order. At any rate, the current debate on materialism focuses on the ontological status of propositional states (such as beliefs), on the one hand, and phenomenal states (such as pains), on the other. Both of these kinds of mental states are considered to occur within the physical world. Hence, what is philosophically puzzling is precisely the relationship between mental events or states, on the one hand, and physical events or states, on the other. Finally, when it comes to the distinction between materialism as a robust doctrine and physicalism as a linguistic doctrine, I find it artificial and motivated largely by an outdated anti-metaphysical spirit. Consequently, I think there is no harm in using the labels “materialism” and “physicalism” interchangeably.

I began by remarking that the overwhelming majority of contemporary philosophers assume that materialism must be true. Let me now come back to this modal claim. This statement of the necessary truth of materialism needs a qualification. Most contemporary materialists would concede, upon reflection, that materialism is a *contingent* thesis. In other words, materialism is not *necessarily* true. Even if it were true about our world, there might be worlds of which it does not hold. Furthermore, many believe that materialism is an *empirical* claim. That is, we do not know *a priori* that materialism is true. Rather, it is a hypothesis that is well confirmed by the previous course of scientific inquiry. To be sure, this leaves a possibility that the future inquiry will go in a direction that will prove materialism false. Yet this possibility appears decreasingly small. Now, it might be that the thesis of physicalism is both empirical and contingent, but we have learned from Kripke that these two notions must not be confused with one another. The former concerns the *epistem-*

⁴ The idea of reductive materialism is expounded here in terms of *nomological* reduction, i.e., the relation between theories, which must be distinguished from an earlier idea of *translational* reduction, i.e., the relation between sentences. The latter is behind the project of logical empiricists, such as Carnap (1932), while the most developed form of the latter is the work of the Australian materialists, such as Armstrong (1969). The classic statement of the token-identity, or nonreductive, materialism is Davidson (1970); an early extension of this approach to other special sciences besides psychology is Fodor (1974).

ic status of physicalism; the latter its *modal* status. Assuming that physicalism is an empirical thesis by no means guarantees that it is contingent. It is now a well-accepted dictum that certain necessary truths could be discovered by empirical means. At any rate, although the thesis that physicalism is only *a posteriori* and contingently true represents by and large a received wisdom, hardly anybody takes seriously the notion that physicalism could have a *rival*.⁵

3. Anti-Individualism

In spite of this near consensus, I think we should take seriously the idea that materialism is false. And when I am saying that we should take this idea seriously, I do not just mean to suggest that materialism *could* be false, though it happens to be true. Rather, I am suggesting that materialism *is* false. To make use of the possible worlds idiom: I claim that materialism is false in the actual world, though there might be worlds in which it is true. Moreover, in supporting this claim, it is not necessary to look for some empirical evidence that counts against the truth of materialism, such as possible loopholes in the current cognitive research. Rather, it is possible to proceed on *a priori* grounds (with a qualification to be explained later in this section). *A priori* refutations of certain research programs are no more seen as ridiculous usurpations of empirical inquiry by antiquated armchair metaphysics, as they had been in the heyday of logical empiricism. Ever since Kripke it is acceptable to maintain that a purely philosophical argument could undermine an empirical thesis. The *a priori* argument that I have in mind here is based on the thesis of *psychological anti-individualism*, in particular in the form developed by Tyler Burge.

In a nutshell, Burge's psychological anti-individualism is a metaphysical theory about the identity conditions of propositional psychological states. According to this theory, beliefs persist or alter under conditions independent of the conditions under which the body of the individual who holds those beliefs persists or alters. Rather than by the body of the

⁵ The crucial distinction between the modal status of physicalism and its epistemic status follows Kripke (1980). Among many of those who hold that physicalism is a contingent claim, see Lewis (1983, 35) and Post (1987, 187). Teller (1983) represents a dissenting minority, claiming that physicalism is necessarily true. Those who see materialism as an empirical thesis include Oppenheim and Putnam (1958), Hellman and Thomson (1977, 311) and Post (1991, 95).

believer, the identity of the beliefs is directly affected by the character of the environment beyond the individual—hence the name “anti-individualism.” The theory of psychological anti-individualism is a result of the familiar Twin Earth thought experiments. Two types of these thought experiments have been proposed. In the first type, environmental changes in question are *physical*. For example, here on Earth, a certain individual, Alf, is ensconced in a physical environment replete with aluminum, and he utters statements such as “Pans and boat masts are made of aluminum.” Given a reasonable assumption that statements express thoughts, Alf’s statement expresses the thought that *pans and boat masts are made of aluminum*. On Twin Earth, Alf finds himself in an environment, which is identical to that of Earth, except that aluminum here is replaced with a metal, which is chemically different from, though phenomenally identical with, aluminum. Call it “twalum” in English. Due to its phenomenal no less than structural identity with aluminum, twalum is put on Twin Earth to the same uses as aluminum. Thus it prompts the same comments from Alf, such as “Pans and boat masts are made of aluminum.” But, since there is no aluminum on Twin Earth, this utterance can only be interpreted as expressing the thought that *pans and boat masts are made of twalum*. The second type of Twin Earth thought experiments assumes no physical, only a *social*, change. Alf on Earth utters numerous statements about arthritis, including some erroneous ones, such as “Arthritis has spread into my thighs.” Given the aforementioned assumption about the connection between beliefs and utterances, should Alf’s statements be interpreted as a belief about arthritis? Yes, since Alf would admit his error as soon as a physician pointed it out to him. Consequently, even prior to a consultation with an expert, Alf on Earth has thoughts about arthritis, including some mistaken beliefs, such as *arthritis has spread into my thighs*. But it so happens that on Twin Earth, all the things that Alf says about his affliction, including “Arthritis has spread into my thighs,” count as correctly applying to the disease called “arthritis.” But, of course, they have no concept *arthritis* on Twin Earth, because arthritis does not affect thighs. For clarity’s sake, call the rheumatoid affliction as conceptualized by Twin Earthians by the English neologism “tharthritis.” Thus, we should say that Alf has no conception, albeit distorted, of arthritis of Twin Earth, but rather the conception of tharthritis. When asserting the statement “Arthritis has spread into my thighs,” he expresses the belief *tharthritis has spread into my thighs*. The common fea-

ture of the two thought experiments is that Alf's body persists unchanged through the variations of thought contents, so that the latter depend on the changing character of environment, be it physical or social.⁶ This already puts a strain on materialism: if, as materialists suppose, the mind is the body, then an individual whose bodily identity is unchanged throughout an environmental change should remain mentally unchanged as well. At this stage, however, we do not have a conclusive anti-materialist argument yet.

Burge's actual argument proceeds under two constraints. First, its scope is limited to propositional states or events—beliefs, desires, and the rest. The argument designed by Burge thus does not affect any claims pertaining to the status of phenomenal states.⁷ Second, the argument targets the token-identity, or nonreductive, form of materialism, briefly introduced in the previous section. As explained there, this theory frames the psychophysical relation in terms of the identity between token beliefs and token physical states or events. As we shall see momentarily, this second constraint does not preclude expanding the argument to cover other forms of materialism, which are defined in terms of relations weaker than identity. At any rate, Burge's argument can be summarized in just two steps. I shall quote Burge's own summary of the argument at length:

The first premiss of the argument against the token-identity theory is strongly suggested (though not entailed) by the thought experiment:

- (1) It is possible for a person to think thoughts with different contents even though all event-tokens that occur in the individual's body, that are plausible candidates for identification with mental events, and that are speci-

⁶ The experiment based on a variation of the social environment was suggested already in Burge (1979). Both the social and physical variations are considered, e.g., in Burge (1986) and numerous later papers. I discussed Burge's two thought experiments in much more detail, including their relationship to somewhat different arguments of Putnam's, in Hříbek (2002).

⁷ To be sure, this is a small consolation to materialists, since there are independent reasons to worry about the place of phenomenal states within the physical world. Indeed, Chalmers (1996) has famously spoke of the "hard problem" of understanding how physical states can possess phenomenal properties, compared to the alleged no-brainer of understanding the possession by some physical states of propositional contents. I am less optimistic than Chalmers as to our ability to already even formulate the problem of phenomenal consciousness; but I disagree even more strongly with his view that a successful naturalization of intentionality is around the corner.

fiable by physical sciences such as physics, chemistry, and neurophysiology, are the same.

The second premiss is less specifically related to the thought experiments:

- (2) No occurrence of a thought could have a different intentional content and be the very same token-event or event-particular.

Now take any physical event-token *b* in the individual's body that is a plausible candidate for being identical with the individual's occurrent thought (mental event-token) *a* that aluminium is a light metal (or that arthritis is a painful disease). By (1), there are possible situations in which the same token *b* occurs, but in which there occur only thoughts (mental event-tokens) with different intentional content. By (2), none of these thought occurrences is the very same token event as *a*. So since *b* could occur without *a*'s occurring, *b* cannot be identical with *a* (Burge 1993, 105).

Clearly, the strategy of Burge's argument consists in presenting a *single counterexample* against the *general thesis* of token-identity theory. In other words, it suffices to show that the belief event-token that *arthritis has spread into my thighs* is not identical with a physical event-token, and the general thesis that every propositional event-token is identical with some physical event-token turns out to be false. But the argument easily generalizes, since it does not matter which example we choose. Therefore, it can be further inferred that *no* thought event is identical with a physical event. In other words, there are some token events that are not physical events – namely, beliefs and other propositional events.

I wish to add a few more words concerning the epistemic and modal character of Burge's anti-materialist stance. He proceeds by considering the actual practice of belief attribution. He assumes no more than the commonplaces that sincere utterances express beliefs, that laymen defer to experts, and such. This way, Burge reveals anti-individualism to be implicit in certain facts about our everyday psychological discourse. Thus the claim that we can learn about the falsity of materialism by *a priori* means is a bit more complicated than stated in the beginning of this section: we have to be familiar with the *actual* character of psychological discourse. On the other hand, it remains true that examining this discourse is certainly unlike conducting scientific experiments. As for the modal status of Burge's claim, it is enough to realize that facts about belief attribution and content fixation *could* be otherwise. So this is why I believe that materialism, though false in our world, could be true if facts about belief identification were different.

4. Three Compatibility Issues

Although psychological anti-individualism is gradually gaining popularity in philosophical circles, its radical consequences for the thesis of materialism have gone either unnoticed or underestimated. Let me distinguish three *compatibility issues* that arise in connection with anti-individualism. A larger part of the literature on anti-individualism is concerned with the question of its compatibility with *privileged self-knowledge*: if the identity conditions of intentional states include aspects of the environment of which the subject of these states might be ignorant, how can she know what she thinks, as she should, according to our deeply rooted intuition?⁸ Another compatibility issue that has been widely discussed has to do with the apparent tension between anti-individualism and *mental causation*: again, if the subject's beliefs are identified with respect to factors beyond the boundaries of the individual subject, how can these beliefs be causes of her behavior, which should be internal to the subject?⁹ Finally, there is the third compatibility issue, which concerns the relationship between anti-individualism and *materialism*. If the individuation conditions of intentional states do depart from those of physical states as shown by Burge's argument, then it appears that mental states cannot be physical states. Why has this last topic been so far neglected, compared to the attention paid to the first two compatibility issues? The reason is that while on the first two issues there are plenty of participants on both sides of the fence, when it comes to the question of the compatibility of anti-individualism with materialism almost everybody is a compatibilist. In other words, almost everybody – including Burge, in some of his moods – seems to be convinced that materialism can be squared with anti-individualism.

⁸ The classic statements of the incompatibility of anti-individualism with first-person authority together with early compatibilist responses are now collected in Ludlow and Martin (1998). Recent work includes Bernecker (2000), Brown (2004), Butler (1996) and Gibbons (1996). There are also two recent collections: Frapolli and Romero (2003) and Nuccetelli (2004). Notice that the question of the compatibility of anti-individualism and privileged self-knowledge is an *epistemic*, not a *metaphysical*, issue.

⁹ The classic statement of the incompatibilist worry that anti-individualistically individuated states cannot be internal causes of behavior is Fodor (1980). See also Block (1986), Churchland and Churchland (1983), Egan (1991), Fodor (1988; 1991), Segal (1989; 2000, chap. 5), and Stich (1983). Compatibilists with respect to this issue include Baker (1995), Burge (1986b; 1989), Dennett (1987c) – but cf. Dennett (1982) – and Millikan (1983; 1993, esp. chaps. 7 and 8).

There have transpired two pairs of proposals that take the threat of anti-individualism seriously, yet aim at reconciling it with one or another version of materialism. One pair of proposals assumes that the relevant relation between the mental and the physical is identity, where the functional realization is understood as a close relative of identity. Since identity is the strongest kind of relation, we can call the corresponding compatibilist project the *strong relation project*.¹⁰ The other pair of proposals involves relations weaker than identity, in particular supervenience and constitution. I propose we call these two versions of the *weak relation project*.¹¹ I believe it can be shown that all these compatibilist projects fail, but I shall not provide detailed refutations in this paper. For one thing, I do not have a space here to argue against any of these projects in a sufficient detail. But, more important, I announced already in section 1 that the purpose of this paper is not to develop a conclusive argument in favor of any particular relation between anti-individualism and materialism, but rather to attempt a basic map of issues and options involved in this relationship. In particular, I am interested in the options that are left given that the two ideas are indeed irreconcilable.

Yet let me at least provide a rough outline of the key reasons for thinking both the strong relation and the weak relation projects are

¹⁰ The representatives of this project are Davidson (1987), Dretske (1996), Harman (1988), MacDonald (1990), Millikan (1993), Pettit (1992), and Sober (1985). Davidson and MacDonald attempt to reconcile anti-individualism with token identity, while Harman, Dretske, Millikan, Pettit and Sober believe in the possibility of squaring anti-individualism with assorted functionalisms.

¹¹ Simply put, supervenience is a relation between kinds of properties. A set of properties *B* supervenes on a set of properties *A* if and only if, for any two objects *x* and *y*, if *x* is identical with *y* with respect to *A*, then *x* is also identical with *y* with respect to *B* (where *A* and *B* are, e.g., physical and mental properties, respectively). Those who wish to achieve a reconciliation between anti-individualism and materialism by means of supervenience usually work *global* supervenience, in which the relevant *x* and *y* are the whole worlds rather than individuals within them. See Beckermann (1992), Heil (1992, chap. 3), MacDonald (1990; 1995), Papineau (1995), Petrie (1987), Seager (1992), and Wilson (1995, chap. 6), etc. Burge himself appears to endorse global supervenience in Burge (1989). Kim, who was the first to suggest the notion of global supervenience in Kim (1982), ultimately rejects it in Kim (1987, 87). In several papers, starting with Burge (1986), Burge also vaguely suggests that the version of materialism that could be reconciled with materialism should be based on the notion of constitution. He never explains what he means by the term “constitution,” but the way it is understood in the current literature, it signifies the relations between material particulars that can coincide in space and time, such as a statues and a piece of marble.

doomed, the reasons that I justify in detail elsewhere.¹² As for the former project, which hopes to square anti-individualism with token identity, it fails because it assumes that thoughts and other propositional states have their contents only contingently. In other words, it assumes that, e.g., the thought that *arthritis has spread into my thighs* is only contingently the thought that *arthritis has spread into my thighs*, since the physical token-event with which it is identified in the actual world persists in other possible situations while the content changes. Yet our practice of belief identification is such that we pick them out by their contents. When it comes to the weak relation projects—those that suggest that relations weaker than identity could suffice for materialism—they appear to compensate for the aforementioned problem of the identity theory by granting mental events a greater autonomy, i.e., to make room for their specific persistence conditions. However, these weaker theories disappoint in another respect. They fail to do justice to some basic intuitions about materialism mentioned in section 2—viz., that the physical, if it is not everything that there is, at least *determines* everything else that there is. So the greater autonomy for mental events is bought at the price of diluting materialism to the point at which it is no longer recognizable as such. All that remains is a vague notion that there is nothing *supernatural* intervenes into the world, but that is, of course, a far cry from any version of materialism that has ever been proposed.

5. Beyond Materialism

So suppose now that materialism and anti-individualism are irreconcilable. Suppose further that anti-individualism is, as suggested at the close of section 3, a reflection on the way we do, as a matter of fact, identify beliefs and other intentional events. It seems to follow that we must keep anti-individualism and give up materialism. Does the demise of materialism leave us with any options?

We seem to be left with three options altogether. The first one is to give up materialism, understood as the thesis that mentalistic vocabulary refers to physical events, for the thesis called, somewhat paradoxically, eliminative materialism, according to which mentalistic words refer to *nothing* at all. Stephen Stich in particular argued for this on the

¹² In chapters 2 and 3 of my unpublished dissertation. See footnote 2 above.

grounds that our everyday ways of identifying propositional events really are anti-individualistic. So anti-individualism is true. But this just shows that these everyday ways are unscientific and should be replaced with something more respectable. The second post-materialist option is to refuse to accept the underlying metaphysical assumption that taxonomies in domains other than physics must be modeled on physical taxonomy. Burge sometimes prefers this option. He argues that metaphysics should wait until we have studied closely enough the many different ways in which explanatory kinds are individuated in different scientific disciplines. The result is a sort of metaphysical pluralism. However, one might argue that such pluralism does not make the question of the metaphysical relation among the explanatory kinds of different sciences illegitimate: the question is only postponed. Finally, we might be pressured to reconsider the third, and final, alternative to materialism, which is dualism. Most philosophers nowadays do not even take this option seriously, although Burge has recently tried hard to present a more plausible version of dualism. However, a critic might easily point out to Burge that his defense of dualism does not even begin to explain how it would deal with the problem of mental causation. This is not to say that mental causation is necessarily a problem for every metaphysics of mind. It is such for Burge's theory, however, because Burge assumes that psychological events are causal antecedents of bodily events.¹³

To be sure, one might hope that the list of alternatives that I have provided is not complete. Perhaps there is position that I have not considered? I agree that a possibility that a new and better alternative will one day be articulated has not been logically ruled out. Even so, I think this possibility gets increasingly remote. This is because I believe that progress in philosophy in general consists in charting a logical geography of some domain: we start with some rough categories and concepts, but over time these are clarified and better distinctions are made, until we reach a point beyond which a seemingly novel proposal turns out to be some old concept under a new name. In listing eliminativism, explanatory pluralism, and dualism as alternatives to all versions of materialism, I have largely followed Burge. But other philosophers, while com-

¹³ For eliminativism, explanatory pluralism, and dualism, see Stich (1983, 165 – 166), Burge (1993), and Burge (2004), respectively. I examine, and reject, these proposals in chapter 4 of my dissertation.

ing from different directions and displaying different commitments, have reached similar results.¹⁴

So I shall just suppose that we cannot expect any groundbreaking conceptual innovation that would expand our list of post-materialist options. In other words, I shall take seriously the idea that the mind-body problem might be unsolvable or intractable. While I have flown quickly through the preceding steps in order to get to the present point, the remainder of the paper will move at a slower pace. As I said already at the outset, most of the claims made in this paper are largely just stated than argued for in detail. However, this is not because I am unable to substantiate them, but rather because I needed them to set the stage for the crucial question: Does the intractability of the mind-body problem spell some kind of intellectual disaster?

Colin McGinn has recently come up with a widely discussed argument that leads to intractability, so I shall use his work as a backdrop in order to explain my own view. McGinn introduces the idea of “cognitive closure,” which he defines as follows:

Cognitive closure. “A type of mind *M* is cognitively closed with respect to a property *P* (or theory *T*) if and only if the concept-forming procedures at *M*’s disposal cannot extend to a grasp of *P* (or an understanding of *T*.” (McGinn 1989, 3)

According to McGinn, human minds are cognitively closed with respect to the properties of the body that explain how it can subserve consciousness. McGinn believes that such properties *P* are bound to exist, unless we are ready to accept either eliminativism or supernaturalism. Also, the requisite theory *T* might not even be exceedingly complex, compared to other theories; after all, consciousness is a property widely instantiated even by relatively simple organisms, which do not exhibit any other higher mental capacities. So, *T* might be simpler than various theories of language and cognition that we already have. It just so happens that our minds are not fit to formulate *T*. Why is that? According to McGinn, the reason is that we conceive of the brain—and, indeed, of the body in gen-

¹⁴ For example, Kim’s logical geography of solutions to the mind-body problem includes reductive materialism and eliminativism, while assuming that dualism is an obvious non-starter. Leaving aside reductive materialism, which is currently Kim’s reluctant choice, this leaves us with eliminativism and dualism. As for nonreductive materialism, such as Davidson’s theory, Kim argues that it collapses into eliminativism. See Kim (1989).

eral—in perceptual terms, whereas consciousness is not a perceptual property of the brain. And since we could never derive the concept of conscious states simply on the basis of what is perceivable about the brain, the former would not be introduced in order to explain the data about the brain, either. So the nature of conscious mental states is cognitively closed to our type of minds.

What is crucial about McGinn's argument is that the nature of the psychophysical relation is non-mysterious and has an explanation in *some* science—it just so happens that this science is inaccessible to *our* kind of mentality. For McGinn, the intractability of the nature of consciousness—and, equally, of the nature of the relation between brains and propositional contents—is a matter of comparative stupidity, or a lack of sensitivity to the presence of some kind of property instantiated in the external world, but accessible only to those with the right kind of intellectual capacity. McGinn's favorite example is a hypothetical Humean mind. This type of mind could be filled only with concepts ("ideas") that copied percepts ("impressions"), and thus it would be unable, for example, to conceive of the unobservable structure of physical objects. In other words, the Humean mind could not build physical science. McGinn speculates that subhuman animals might have Humean minds. Be that as it may, *we* don't have such minds, and so the structure of physical reality is cognitively open to us. But perhaps our own type of mind is cognitively closed to metaphysical problems, such as the nature of consciousness—or, closer to the topic I have tackled in my essay, the nature of the relationship between brains and propositional contents. Yet this leaves out the possibility that yet another type of intelligence—either perceptive to the facts that allude us, or more logically astute—would be capable of solving such metaphysical mysteries (see McGinn 1989, 4).

By contrast, I wish to argue that the apparent insolubility of the mind-body problem is not just a matter of our stupidity. Rather, it is the matter of a tension between mutually irreconcilable conceptual commitments—a tension that is insoluble, period, not just insoluble by us. Let me set the stage by presenting an argument that seems to suggest something similar to what I have in mind. In his well-known puzzle about belief, Kripke offers the following thought experiment (see Kripke 1988). Pierre grew up in France as a monolingual speaker of French surrounded by other such monolingual Frenchmen. In his childhood, his nanny used to tell him stories about a beautiful distant city called in

French “Londres.” On the basis of the stories by his nanny, Pierre formed a belief that this foreign city was beautiful, and he expressed this belief by saying: “Londres est jolie.” As an adult, Pierre moved to London, England, and he learned English without any access to a translation manual; instead, he picked up the local language from the English-speaking natives that surrounded him. Also, Pierre unfortunately settled in an ugly part of the city. So he is disposed to say, in his newly acquired English, “London is not pretty.” However, what Pierre does not know is that the city he had learnt about in France is the very city he now resides in. Now suppose that Pierre’s logical capacities are impeccable, so he would not entertain contradictory beliefs. And yet he must be attributed a contradictory belief—viz., the belief that London is and is not pretty. This is precisely the “puzzle about belief,” because the attribution of an inconsistent belief to Pierre wreaks havoc of our normal practice of belief attribution. It is difficult to say what message exactly is Kripke trying to get across, but on one interpretation, he seems to be suggesting that our practice involves a commitment to various principles, each of which appears fine on its own, but it is inconsistent with all the rest. Thus, for example, there does not seem to be anything unacceptable about the assumption—Kripke calls it “the principle of disquotation”—that sincere utterances express the contents of the speaker’s mind. Accordingly, it seems fine to assume that when Pierre assents to “Londres est jolie,” he expresses his belief that London is pretty. Further, there appears to be nothing problematic about the principle that translation preserves reference, so that when Pierre assents to “Londres est jolie” and “London is not pretty,” he has beliefs about one and the same city, because both “Londres” and “London” refer to London. Finally, although perhaps more controversially, it is assumed that a certain anti-individualist view of content determination holds, since there is simply just one city in the external world that Pierre’s various beliefs refer to, and which thus determines their contents.¹⁵ As a result of following these seemingly fine principles, however, we end up attributing an inconsistent belief. This implies, however, that our psychological discourse is torn by incon-

¹⁵ The anti-individualist view of reference is not so explicitly distinguished by Kripke himself as one of the commitments that we bring in to our practice of belief attribution. However, given that Kripke spends some time early in his paper by reviewing his criticism of description theories of reference, we can assume that the causal picture of reference is in the background of his discussion of the “puzzle.”

sistency. And, as Kripke freely admits, there does not seem to be any way out of this situation.¹⁶

Kripke's claim that impeccable commitments can result in the attribution of an inconsistent belief and that, by implication, our trusted everyday practice of intentional discourse can seem to be incoherent, is similar to my point that the mind-body problem might be unsolvable in principle. We seem to have a good reason to believe that the identity conditions of attitudes that we attribute to each other are anti-individualistic, or externalistic. At the same time, we assume that attitudes are entertained by embodied subjects in the ultimately physical world. Many philosophers—including Burge, as noted earlier in this section—wish to unpack this latter commitment in the form of some strong causal conception of psychological explanation, which construes beliefs and other types of intentional events as antecedent causes of behavior. But, even if such a strong interpretation proves unwarranted, we seem to remain committed to regarding mental states and events as interacting with physical states and events. At any rate, none of these commitments seems to be something that could be overthrown by piling up more facts about psychology, including facts that might forever be inaccessible to us. The aforementioned commitments can be discovered by a mere reflection on our everyday practice. Yet put together, anti-individualism does not cohere with the notion that we are embodied thinkers. We are pulled toward various versions of materialism, on the one hand, and various attempts to abandon materialism, on the other. But we seem to have run out, or are running out, of substantive options. So it does not seem that we could be able to conceive of the mind-body relation that respected all our commitments, only if we knew some facts that are presently inaccessible to us, or if we were more logically apt.

Let me conclude by suggesting that the situation that transpires in consequence of our inability to solve the mind-body problem parallels our predicament with respect to the problem of induction. I mean the problem that Hume formulated more than two hundred and fifty years

¹⁶ Kripke considers four possible solutions of the paradox: (1) Pierre did not believe that London was pretty when he was a monolingual Frenchman. (2) Pierre did believe that London was pretty in his monolingual French past, but he does not believe that London is not pretty now, when he is a bilingual resident of London. (3) He neither believed that London was pretty in the past, nor does she believe that London is not pretty now. (4) He believes both. Kripke shows that none of these avenues is credible. See Kripke (1988, esp. 120 – 123).

ago, and no one has positively solved ever since. The Humean problem of induction is, of course, the problem of justifying our beliefs concerning unobserved events, including future events. Every day, we tend to reason: The *F*s that I have experienced so far have been *G*s. Therefore, the next *F* that I shall experience will also be *G*. Or even: All *F*s are *G*s. The question is, how do we justify this general conclusion? It is clear that it does not deductively follow from the premise, which is a statement of empirical fact. In order to form a valid argument, we must spell out a hidden premise that the argument relies upon: If a regularity *R*—e.g., that All *F*s are *G*s—holds in my experience, then it holds in nature in general. In other words, whenever we reason inductively, we assume that our experience is representative of the rest of nature. But this assumption about *R* is in itself a claim about unobserved events, hence it would have to be justified inductively. But this claim is a premise in any inductive argument. So any inductive argument for this claim would be circular. So there is no positive resolution of the problem of induction. Does this mean that we are being irrational every time we reason inductively? We believe of ourselves that we are rational beings, and though we have our moments, we do make genuine attempts to eliminate irrationality from our lives. But given that inductive reasoning fills out most of our waking lives, it is inconceivable what it would be like to eliminate it. Fortunately, such a drastic change of attitude is not only impossible; it is not mandated by our commitment to rationality, either. For Hume suggests that our unjustifiable generalization of *R* is not irrational. Though it is not logically compelling, it is not inconsistent, either. Moreover, it is natural for creatures like us, in the sense that we cannot help using it. I think our commitment to intentional vocabulary, to seeing ourselves and our fellow human beings as intentional agents acting on beliefs, is similarly part of our nature. If the reasoning on the preceding pages is sound, we might not be able to come up with a coherent picture of the physical world that includes minded creatures in it, but it does not follow that we should try to do the impossible and fashion ourselves either as mere bodies (as recommended by eliminative materialism) or, alternatively, as disembodied souls (as implied by dualism). Our self-image as intentional agents is as unproblematic as our view of ourselves as rational inquirers. Our inability to formulate a satisfying metaphysics of mind may be less a proof of the inferiority of our rationality than of the incoherence among its constituent claims.

Filosofický ústav AV ČR
Jilská 1
110 00 Praha 1
tomas_hrbek@hotmail.com

REFERENCES

- ARMSTRONG, D. (1969): *A Materialist Theory of the Mind*. London: Routledge.
- BAKER, L. R. (1987): *Saving Belief*. Princeton: Princeton University Press.
- BECKERMANN, A. (1992): Supervenience, Emergence and Reduction. In: Beckermann, A. – Flohr, H. – Kim, J. (eds.) (1992): *Supervenience or Reduction? Essays on the Prospects of Nonreductive Materialism*. Berlin: de Gruyter, 94 – 118.
- BERNECKER, S. (2000): Knowing the World by Knowing One's Mind. *Synthese* 123, 1 – 34.
- BLOCK, N. (1986): Advertisement for a Semantics of Psychology. In: French, P. – Uehling, T. – Wettstein, H. (eds.): *Midwest Studies in Philosophy* 10. Minneapolis, MN: University of Minnesota Press, 615 – 678.
- BROWN, J. (2004). *Anti-Individualism and Knowledge*. Cambridge, Mass.: The MIT Press.
- BURGE, T. (1979): Individualism and the Mental. In: French, P. – Uehling, T. H. – Wettstein, H. K. (eds.): *Midwest Studies in Philosophy* 4. Minneapolis, MN: University of Minnesota Press, 73 – 121.
- BURGE, T. (1986): Individualism and Psychology. *Philosophical Review* 45, 3 – 45.
- BURGE, T. (1989): Individuation and Causation in Psychology. *Pacific Philosophical Quarterly* 70, 303 – 322.
- BURGE, T. (1993). Mind-Body Causation and Explanatory Practice. In: Heil – Mele (1993), 97 – 120.
- BURGE, T. (2003): Descartes, Bare Concepts, and Anti-Individualism: Reply to Normore. In: Hahn, M. – Ramberg, B. (eds.) (2003): *Reflections and Replies: Essays on the Philosophy of Tyler Burge*. Cambridge, Mass.: The MIT Press, 291 – 334.
- BUTLER, K. (1997): Externalism, Internalism, and Knowledge of Content. *Philosophy and Phenomenological Research* 57, 773 – 800.
- CARNAP, R. (1932): Psychology in Physical Language. *Erkenntnis* 3, 107 – 142. Reprinted in: Ayer, A. J. (ed.): *Logical Positivism*. New York: Free Press, 1966, 165 – 198.
- CHALMERS, D. (1996). *The Conscious Mind*. Oxford: Oxford University Press.
- CHURCHLAND, P. M. – CHURCHLAND, P. S. (1983): Stalking the Wild Epistemic Engine. *Noûs* 17, 5 – 18.
- DAVIDSON, D. (1970). Mental Events. In: *Essays on Actions and Events*. Oxford: Clarendon, 207 – 225.
- DAVIDSON, D. (1987): Knowing One's Own Mind. *Proceedings and Addresses of the American Philosophical Association* 60, 441 – 458.

- DENNETT, D. (1982): Beyond Belief. In: Dennett (1987), 117 – 202.
- DENNETT, D. (1987): *The Intentional Stance*. Cambridge, Mass.: The MIT Press.
- DRETSKE, F (1996): *Naturalizing the Mind*. Cambridge, Mass.: The MIT Press.
- EGAN, F. (1991): Must Psychology Be Individualistic? *Philosophical Review* 100, 179 – 203.
- FODOR, J. (1974): Special Sciences. Reprinted in: Fodor (1981).
- FODOR, J. (1980): Methodological Solipsism as a Research Strategy in Cognitive Psychology. Reprinted in: Fodor (1981), 225 – 253.
- FODOR, J. (1981): *Representations*. Cambridge, Mass.: The MIT Press.
- FODOR, J. (1988): Making Mind Matter More. *Philosophical Topics* 17, 59 – 79.
- FRAPOLLI, M. – ROMERO, E. (eds.) (2003): *Meaning, Basic Self-Knowledge, and Mind: Essays on Tyler Burge*. Stanford: CSLI.
- GIBBONS, J. (1996): Externalism and Knowledge of Content. *Philosophical Review* 105, 287 – 310.
- HARMAN, G. (1988): Wide Functionalism. In: Schiffer, S. – Steele, S. (eds.): *Cognition and Representation*. Boulder, CO: Westview Press, 1988, 11 – 20.
- HEIL, J. (1992): *The Nature of True Minds*. Cambridge: Cambridge University Press.
- HEIL, J. – MELE, A. (eds.) (1993): *Mental Causation*. Oxford: Oxford University Press.
- HELLMAN, G. – THOMPSON, F. (1977): Physicalist Materialism. *Noûs* 11, 309 – 345.
- HŘÍBEK, T. (2002): Externalismus, aneb vytrvalost karteziánismu. *Filosofický časopis* 50, 359 – 386.
- KIM, J. (1982): Psychophysical Supervenience. *Philosophical Studies* 41, 51 – 70. Reprinted in: Kim (1993), 175 – 193.
- KIM, J. (1987): “Strong” and “Global” Supervenience Revisited. *Philosophy and Phenomenological Research* 48, 315 – 326. Reprinted in: Kim (1993), 79 – 91.
- KIM, J. (1993): *Supervenience and Mind*. Cambridge: Cambridge University Press.
- KIM, J. (1998): *Mind in a Physical World*. Cambridge, Mass.: The MIT Press.
- KRIPKE, S. (1980). *Naming and Necessity*, Cambridge, Mass.: Harvard University Press.
- KRIPKE, S. (1988): A Puzzle About Belief. In: Salmon, N. – Soames, S. (eds.): *Propositions and Attitudes*. Oxford: Oxford University Press, 102 – 148.
- LEWIS, D. (1983): *Philosophical Papers*. Vol. I. Oxford: Oxford University Press.
- LOAR, B. (1988): Social Content and Psychological Content. In: Grimm, R. H. – Merrill, D. D. (eds.): *Contents of Thought*. Tucson: University of Arizona Press, 99 – 110.
- LUDLOW, P. – MARTIN, N. (eds.) (1998): *Externalism and Self-Knowledge*. Stanford: CSLI.
- MACDONALD, C. (1990): Weak Externalism and Mind-Body Identity. *Mind* 99, 387 – 404.
- MACDONALD, C. (1995): Psychophysical Supervenience, Dependency, and Reduction. In: Savellos – Yalçın (1995), 140 – 157.

- MCGINN, C. (1989): Can We Solve the Mind-Body Problem? *Mind* 98, 349 – 366.
Reprinted in: McGinn, C. (1991): *The Problem of Consciousness*. Oxford: Oxford University Press, 1 – 22.
- MILLIKAN, R. (1984): *Language, Thought, and Other Biological Categories*. Cambridge, Mass.: The MIT Press.
- MILLIKAN, R. (1993): *White Queen Psychology and Other Essays for Alice*. Cambridge, Mass.: The MIT Press.
- NEURATH, O. (1931): Physicalism. In: Cohen, R. S. – Neurath, M. (eds.): *Philosophical Papers, 1913-1946*. Dordrecht: Reidel, 1983, 52 – 57.
- NUCCETELLI, S. (ed.) (2003): *New Essays on Semantic Externalism and Self-Knowledge*. Cambridge, Mass.: The MIT Press.
- OPPENHEIM, P. – PUTNAM, H. (1958): Unity of Science as a Working Hypothesis. In: Feigl, H. et al. (eds.): *Minnesota Studies in the Philosophy of Science* 2. Minneapolis: University of Minnesota Press.
- PAPINEAU, D. (1995): Arguments for Supervenience and Physical Realization. In: Savellos – Yalçın (1995), 226 – 243.
- PETRIE, G. (1987): Global Supervenience and Reduction. *Philosophy and Phenomenological Research* 48, 119 – 130.
- PETTIT, P. (1992): *The Common Mind: From the Intentional Psychology to Social and Political Theory*. Oxford: Oxford University Press.
- POLAND, J. (1994): *Physicalism: The Philosophical Foundations*. Oxford: Oxford University Press.
- POST, J. (1987): *Faces of Existence: An Essay in Nonreductive Metaphysics*. Ithaca, NJ: Cornell University Press.
- POST, J. (1995): "Global" Supervenient Determination: Too Permissive? In: Savellos – Yalçın (1995), 73 – 100.
- ROBINSON, H. (1993): Introduction. In: Robinson, H. (ed.): *Objections to Physicalism*. Oxford: Clarendon.
- SAVELLOS, E. E. – YALÇIN, Ü. D. (eds.) (1995): *Supervenience: New Essays*. Cambridge: Cambridge University Press.
- SEAGER, W. (1992): Externalism and Token Identity. *Philosophical Quarterly* 42, 439 – 448.
- SOBER, E. (1985): Putting the Function Back into Functionalism. *Synthese* 64, 165 – 193.
- STICH, S. (1983). *From Folk Psychology to Cognitive Science*. Cambridge, Mass.: The MIT Press.
- WILSON, R. A. (1995): *Cartesian Psychology and Physical Minds*. Cambridge: Cambridge University Press.