Article
Eliminative Materialism: Going Concern or Passing Fancy?

CRISPIN WRIGHT

1. A Concessive Form of Eliminative Materialism

Will the recent outbreak of respect—even support—for eliminative materialism (EM) be viewed by our successors as a late twentieth-century aberration or the beginning of wisdom? Barbara Hannan's helpful overview of the debate makes a case for the former assessment. Her argument divides into two broad points. On the one hand she contends that, insofar as there are clear positive arguments for EM, they entirely depend on the conception of ordinary commonsense intentional psychology as incorporating a rudimentary *empirical theory* which, there is purported reason to think, will be superseded by mature science—and superseded not just in the detail of its hypotheses, but in the basic concepts and categories of explanation which it distinctively employs. But this view, Hannan argues— the 'theory-theory' of commonsense psychology (CP)—can and should be rejected. Second, Hannan emphasises that, notwithstanding the questionable character of certain standard attempts to demonstrate that EM is self-defeating, it remains that eliminative materialists have given us not even the most rudimentary explanation of how it might be practical to dispense with the concepts of CP, nor the slightest inkling of a way of thinking about the province of (what we now conceive of as) human rational agency which might supplant them. Ordinary intentional psychology, Hannan reminds us, cannot be entirely wide of the mark if it is granted that there is legitimate use for the concepts of cognition and rational action, for the applicability of CP's distinctive modes of explanation is presupposed in

Address for correspondence: Department of Logic and Metaphysics, University of St Andrews, St Andrews, Fife KY16 9AL. Email: cjgw@st-andrews.ac.uk.

those notions. And the fact is that we are simply bereft of any concrete idea of what it would be altogether to dispense with those notions and their cognates in all our thinking about human behaviour.

My concern here will be not with the correctness of the considerations which Hannan marshals, but with how far they take the friend of CP. What basically drives the eliminative materialist is the suspicion that nothing real answers to the distinctive concepts of propositional attitude psychology—that there are really *no such things* as the kinds of state—beliefs, hopes, desires, and so on—which CP configures in its explanations. The characteristic response of EM to that suspicion is—as the name suggests—to canvass the eradication of CP as the elimination of a super-stition. But that is not, of course, the only possible reaction. The conviction that a discourse deals in no real states of affairs may—in cases where we cannot see our way clear to dropping it, or where it has some special importance or utility for us—provoke not an eliminativist but an *instrumentalist*, or *expressivist*, response. The conviction, for example, that there are no real moral states of affairs can provoke the idea, argued by writers such as John Mackie (1977), that ordinary moral thought is guilty of a sweeping, systematic mistake; but it may also provoke the more conciliatory idea that ordinary moral thinking is not in the business of aiming at truth, realistically conceived, and that we do better to seek a non-factualist construal of moral argument and moral discourse in general.

So in the present case. If Hannan is right that there is no clear prospect of dispensing with the distinctive concepts of CP, that is no more an argument for realism about it than an inability to dispense with evaluative ethical thought is an argument for realism about ethics. And if, in addition, she is right to question the 'theory-theory' of CP, is not that, in effect, to query whether CP deals in the *real causal antecedents* of (much of) human behaviour—and hence, since explanation is manifestly what it purports to offer, to raise a question, entirely congenial to the irrealist spirit which informs EM, about whether CP deals in anything real at all? In short, both of Hannan's principal points: that CP is not an empirical theory, nor—at least foreseeably—a scheme of concepts which we could abandon, are entirely consistent with and congenial to a concessive form of eliminative materialism—one which is content to view the real as exhaustively material, and to 'eliminate' CP not altogether but only from the realm of discourses whose pretension is to describe what is real.

This concessive form of EM seems to be untouched by Hannan's paper. And two forms of motivation for it, each likewise unaddressed by Hannan, may be envisaged. The first, in effect already noted, is best couched by way of an explicit dilemma. Let it be granted that the world is, ultimately, exhaustively material (whatever that means). Commonsense psychology purports to explain certain material goings-on. Are we to understand it as presenting the 'explanatory' states to which it adverts as the real material antecedents of these goings-on or not? If so, then the 'theory-theory' view is after all correct, and the lines of argument which purport to highlight

the deficiencies of CP, so conceived[1]—lines of argument whose sting Hannan tries to draw by questioning the 'theory-theory' view—must be allowed whatever destructive force they have on that assumption. (It is another question how much destructive force that is.) If, on the other hand, it is urged that CP is *not* in the business of depicting the real material antecedents of its explananda, then, since it is unquestionably in the business of explanation of *some* sort, that seems tantamount to the concession that it does not deal in reality at all—that it is not in the business of saying what really produces what—but that its stock-in-trade is rather an 'heuristic perspective', or 'useful stance', whose point and value must be dissociated from the question of application to the real world.

This first line of motivation depends firmly on an unargued materialism as a premise. The second line is more subtle. If token propositional attitudes really exist, then they exist as individuated under a *type*. A type propositional attitude may be conceived as a pair consisting of an attitudinal component—belief, desire or whatever—and the content given by the that-clause used in its specification. So a belief in the reality of the states posited by CP is thus implicitly a belief in the reality of content. An irrealism about psychological content, whatever its detail, must accordingly enjoin a corresponding irrealism about CP.

Arguments which, one way or another, call into question the reality of content, are familiar from the writings of Quine, Kripke, Putnam and others, and number among the most striking products—certainly the most widely discussed—of mid to late twentieth century analytical philosophy. If any argument of this broad category is sound, and if its force can be accommodated, rather along the lines of Kripke's well known 'Sceptical Solution', by conceding that there are no real facts concerning content while conserving the integrity of content discourse by reserving for it some other, non-fact-stating role, then the concessive form of EM will, in effect, be enforced.

2. Boghossian's Argument

I shall not here attempt to review any of the extant argumentation for irrealism about content. What I want to consider in the remainder of this note is a striking recent argument which, if successful, will offer a decisive resource to the opponent of the concessive form of EM.

The argument in question is propounded by Paul Boghossian (1990,

---

[1] Hannan mentions the considerations that CP fails to explain such phenomena as sleep, dreams, mental illness and the behaviour of brain-injured subjects, that CP is unlikely to reduce smoothly to physical theory and thus to participate in a unified science, and that computational psychology is already scoring explanatory successes working with internal states which are not individuated by content.

pp. 175–6), and its gist is that any viewpoint which would deny the factuality of content while granting that discourse about content retains a form of integrity under a non-fact-stating construal, is actually incoherent. Boghossian assumes[2]—and I agree—that there can be no stable marriage of non-factualism about psychological content with factualism about the content of linguistic expressions, so that one who would follow the concessive EM line, granting the legitimacy but denying the factuality of the claims about mental content which are the lifeblood of CP, must take essentially the same tack when it comes to attributions of meaning—and in particular truth-conditions—to linguistic expressions. Such a view, Boghossian argues, can offer no consistent account of the nature of the truth-predicate. Why?

According to Boghossian (1990, p. 161):

> What all non-factualist conceptions have in common—what in effect is constitutive of such a conception of a declarative sentence of the form 'x is P'—is

(1) The claim that the predicate 'P' does not denote a property

—because it is no part of its semantic function so to do—

and hence

(2) The claim that the overall (atomic) declarative sentence in which it appears does not express a truth-condition.

If this is an acceptable formulation of the heart of such views, then it follows that the non-factualist about linguistic content (henceforward 'the content non-factualist') will endorse each of the following.[3]

(5) The predicate, 'has the truth-condition that P', does not refer to a property;

(6) For any sentence S and propositional content P, 'S has the truth-condition that P' is not a truth-conditional claim.

And the striking thing about these contentions, Boghossian argues,—in contrast to the situation of the corresponding claims ingredient in non-factualism about other subject matters—is that they actually contradict each other! For (5) entails that

(7) 'True' does not refer to a property,

whereas (6), he maintains, entails the opposite.

How is that? Why, to begin with, should it be supposed that (5) entails that 'true' does not refer to a property? Boghossian cites his indebtedness to me for the (putative) reason.[4] It is an application of the very vague thought that 'non-factuality' among the determinants of an issue must divest the issue of factuality, and the reflection that S's truth-condition, if it has one, is among the determinants of its truth value. Recast for our present concerns, the question is whether to accept the principle that a predicate ('is true') cannot 'denote a property'—cannot serve to ascribe a real characteristic—if its applying to a particular item is constituted, in part, by another predicate's ('has the truth-condition that P') applying to it which itself is viewed as denoting no real property but as discharging a semantic function of a quite different kind. Doubtless the principle would profit from detailed scrutiny. But it is plausible enough: if sentences' truth-conditions are no real characteristics of them, then we can, as it were, fix all the real characteristics of items in the world without determining what truth-conditions sentences have. But then their truth values, which depend on an assignment of truth-conditions, will be indeterminate as well. Since all the real characteristics are, by hypothesis, fixed, it follows that truth values do not belong in that category.

What about the other half? Why does Boghossian think that (6) entails that 'true' *does* refer to a property? The intuitive train of thought here is that since sentences of the form, 'S has the truth-condition that P', are significant declarative sentences, the claim that they are not in the market for truth and falsity demands something more than a *deflationary* notion of truth (since any significant declarative sentence is apt for deflationary truth and falsity). And the hallmark of all non-deflationary conceptions of truth—the respect, indeed, in which they contrast with deflationism—is their taking 'true' to refer to a substantial property of statements, a property which a *correctly assertible* statement can lack.

Might the effect of this ingenious juxtaposition be resisted by denying that the content non-factualist must accept (5) and (6)? I don't see how. Suppose we use 'true' as a *substantial* truth-predicate, restricted to discourses which satisfy whatever the constraints are which signal genuine 'factuality'; and let us reserve 'correct' as a *minimal* (deflationary) truth-predicate, for which the sentences of a discourse are in the market merely

---

4   Wright, 1984, p. 769.

in virtue of their satisfaction of certain syntactic markers of assertoric content—for instance, the capacity to serve without grammatical incongruity as the antecedent of a conditional—and their subservience to some notion of proper assertibility. Then the content non-factualist's commitment to (6) is immediate; his thesis is precisely that content-discourse in general, and hence ascriptions of truth-conditions in particular, have at best correctness-conditions and lack truth-conditions. As for why content non-factualism should accept (5), that predicates of the form, 'has the truth-condition that P', refer to no property—the reason is simply that corresponding to the distinction between truth-apt and merely correctness-apt discourse will be a distinction between two types of predication, one associated with the generation of statements possessed of substantial truth-conditions—substantial constraints to be met or not by states of the real world—and the other generating only correctness-apt statements. The claim that a predicate 'refers to a property' is then to be understood as placing it in the former group. When (5) is so construed, the content non-factualist's commitment to it follows directly from the reflection that the ascription of a truth-condition is the ascription of a content.

3. How Concessive EM Can Slip the Net

It seems to me that there is nothing importantly amiss either with Boghossian's characterisation of content non-factualism or with the two ingredient lemmas of his argument. How, then, might it be attacked? The point of vulnerability is at the *connection* of the two lemmas. It needs to be clear that the conclusion of the first is indeed inconsistent with the conclusion of the second. But there is a strong suspicion that the argument works only by equivocation on 'refers to a property'. The fact is that the sense in which, according to (5), 'true' *fails to refer* to a property is that predications of it fail to generate contexts with substantial truth-conditions. But the sense in which 'true' *does refer* to a property is that, by (6)—and indeed on any view which allows some correctly assertible sentences not to be truth-apt—the truth-predicate can allow of no uniformly deflationary construal and therefore presumably marks some sort of real representational achievement, as it were, on the part of its bearers. So we have that 'S is true' is not truth-conditional, and that 'is true' is not deflationary. Maybe there is some sort of tension between these claims; but can we elicit an explicit inconsistency?

Yes, provided we make a further (natural enough) assumption. This is that the 'real representational achievement' of any non-deflationarily true statement has to be something which may be marked by a (further) *true* statement, that the fact of a statement's representation of a real matter of fact must itself be a real matter of fact. Granted that assumption, then 'S is true' will be a statement-form whose instances are themselves in the market for truth rather than mere correctness. By the same token a state-

ment's *failure* to represent a real matter of fact, whether resulting from falsity or from its mere correctness, will likewise be a real matter of fact. So the further assumption, generalised, is that the taxonomy of the distinction, to which any non-factualist is committed, between merely correct statements and those which are substantially true, must itself be a *substantial* matter—something to be described by substantially true statements. There is then a direct contradiction with Boghossian's first lemma—effectively that, given (5), instances of 'S is true' are invariably at best correct.

Boghossian's argument will succeed, then, if—and so far as I can see, only if—the content non-factualist can be shown to be committed to the described assumption. But a moment's reflection shows, to the contrary, that he is actually independently committed to rejecting it. If the taxonomy of the distinction between the true and the correct is to be described by true statements, then the same holds for the distinction between discourses whose best sentences, as it were, are true and those whose best sentences are merely correct. But that is the distinction between factual and non-factual discourses. So the assumption entails that statements which classify discourses on one side or the other of the factual/non-factual distinction must themselves be apt for truth rather than correctness, that the details of that distinction must be factual. And now it is obvious that the assumption is one with which the content non-factualist can have no truck, since how a discourse fares under the distinction will be a function of the species of *content* its statements possess, and therefore—by the very principle which informs the argument from (5) to (7)—*cannot* be a factual matter.

In sum: the two lemmas of Boghossian's argument are in genuine contradiction only if we take it that the very existence of a distinction between the truth-apt and the merely correctness-apt demands that its details may be recorded by *truths*. But Boghossian seemingly fails to notice the role of this assumption in his argument, let alone argue for it. Why? Apparently because he takes it that there is, across all discourses, *only one truth-concept around*—so that the taxonomy of *any* distinction can only be recorded by statements which are true by its lights. This monism about truth is explicit in the following passage (Boghossian, 1990, p. 165, footnote 17):

> Whether truth is robust or deflationary constitutes the biggest decision a theorist of truth must make. But decide he must. It is an assumption of the present paper that the concept of truth is *univocal* as between these two conceptions, that a concurrent commitment to *both* a robust and a deflationary concept of truth would be merely to pun on the word 'truth'. We should not confuse the fact that it is now an open question whether truth is robust or deflationary for the claim that it can be both. There is no discernible plausibility in the suggestion that the concept of a correspondence between language and the world and the concept

of a language-bound operator of semantic ascent might both be versions of the same idea.

It is remarkable that Boghossian here helps himself to a thesis which is already in tension with *any* form of non-factualism, whether about content or anything else. For the non-factualist can hardly deny that a minimal (deflationary) truth-predicate will be definable over the sentences of his target discourse—sentences which he holds to be ineligible for the more substantial property associated with 'genuine factuality'. And Boghossian's monism acquires in any case no support from the implausibility of the idea that a single concept of truth might somehow incorporate both correspondence and deflationary ('language bound') aspects. That is not the question. Nor is the question whether distinct concepts of truth might operate over a single discourse. The question is whether distinct truth-predicates may be in play over *distinct* discourses, whether 'true' as applied to moral judgements, for example, may mark a different characteristic to that which it signifies when predicated of judgements concerning the primary qualities of material bodies. Boghossian's discussion contains no cogent objection to this suggestion.

This is the crux of the matter. It ought to have been clear in advance, independently of the detail of Boghossian's argument, that if the content non-factualist has only the resources of a *single* notion—undifferentiated 'truth'—by means of which to express his views, he will have to allow that the existence of a distinction between 'non-factual' and 'factual' assertoric discourses—however he proposes to characterise it—will depend on something's being *true* of the latter which is not true of the former. If he then goes on to say that the distinction consists in the fact that the statements of 'factual' discourses possess truth-conditions, whereas those of 'non-factual' discourses do not, he will have committed himself to the view that the distinction is constituted in a province—content—about which, according to his own view, there is nothing *true* to say! In that case he will indeed have tied his own shoelaces together (if Boghossian has not already tied them for him).

The moral, Boghossian's plea for the univocity of 'true' notwithstanding, is to recognise that any content non-factualist *must* work with a pluralist conception of truth. In particular, he has to insist that the distinction between the truth-apt and the merely correctness-apt is one whose details, since founded in semantic contrasts, may be recorded only by *correct* statements. 'S is true', 'S has the truth-condition that P', 'S is correct', 'S has the correctness-condition that P'—all these will be statement-forms whose instances are, for the content non-factualist, only correctness-apt. Once that is clear, he may cheerfully and consistently grant both that 'true' does not refer to a property—meaning thereby just that instances of 'S is true' are only correctness-apt—and that 'true' does refer to a property,

thereby signalling his repudiation of a globally deflationist account of truth.[5]

4. Conclusion

I do not mean to befriend content non-factualism. There *are*, of course, serious potential difficulties for the position. For one, there is the suspicion, hard to clinch, that it must inflate into a global non-factualism and that the latter is incoherent.[6] For another, suppose that the content non-factualist avoids Boghossian's contradiction in the way I have suggested. Then he will have to grant that the question, whether content non-factualism is to be accepted, is itself, like all questions to do with the demarcation of the truth-apt from the merely correctness-apt, a question apt only for a *correct* answer, rather than a substantially true one. But it seems likely, on any plausible account of the distinction between truth and correctness, that merely correct opinion will turn out to contrast with true opinion by lacking a certain kind of *cogency*—in the way that, say, opinions about what is funny lack the cogency of opinions about what is cubic.[7] And if so, then content non-factualism must itself lack that kind of cogency, however impressive the arguments used in its support.

We may stand to inflict some damage on content non-factualism, therefore, by following through the implications of the way I have outlined whereby it can resist Boghossian's attack. But there is, unfortunately, no simple contradiction inherent in the view along the lines Boghossian tried

---

[5]   It might seem that the content non-factualist has another option, namely to retain a monistic view of truth and seek to characterise the distinction between non-factual and factual discourses in quite other terms than in point of truth-conditionality. But this is not really another option (provided the alternative characterisation locates the distinction in *some* sort of contrast in species of content). Will non-factual discourses turn out to be truth-apt, on the alternative proposal, or will they not? If they do not, then since it remains that they will be subject to *some* notion of stable correct assertibility, the theorist *will be* recognising a plurality of truth-predicates—his monism will be merely verbal. If, on the other hand, non-factual discourses will be truth-apt on the alternative proposal, then however that alternative goes, there can be no good objection to a terminological revision whereby the word, 'true', is annexed to the combination of truth, monistically understood, plus factuality, while 'correct' is reserved for the combination, truth plus non-factuality. And then the point will stand, for just the reason outlined in the text, that the taxonomy of the distinction between the non-factual and the factual cannot, on pain of incoherence, be supposed by the content non-factualist to be capable of more than *correct* statement. In general, so long as the distinction between the non-factual and the factual is grounded in species-content, the content non-factualist will be forced to regard the taxonomy of the distinction as a non-factual matter.

[6]   See Wright, 1984, pp. 769–70. But see also Boghossian, 1989, p. 525.

[7]   For exploration of the link between cogency and more-than-minimal truth, see Chapters 3 and 4 on 'cognitive command' of Wright, 1993. See also Chapter 6 for detailed discussion of the transition from content non-factualism to global non-factualism.

to describe,[8] and no other extant argument against it promises so clean a kill. We are accordingly in no position to conclude that the concessive form of EM, when motivated by doubts about the reality of psychological content, cannot live with the inevitable extension of its view to linguistic content. Nor indeed, in my opinion, are we in any position to affirm that the best of the extant content-sceptical arguments do not impose some form of anti-realism about content. The shift to the concessive form of EM is therefore not merely a way whereby the materialist may avoid the labour of contesting the considerations rehearsed by Hannan; it is something which may yet prove to be imposed on us all. And the question, like so many of the fundamental issues in the philosophy of mind, crosses the boundary into the philosophy of language and the metaphysics of meaning.[9]

Department of Logic and Metaphysics
University of St Andrews
Scotland KY16 9AL
UK

Department of Philosophy
University of Michigan
Ann Arbor, MI 48109
USA

---

[8] Although Boghossian charges that the content non-factualist is committed to contradictory claims about the predicate, 'true', we can represent his basic thought more simply if we focus on the predicate, '. . . has the truth-condition that P'. The thesis of content non-factualism may be represented as

(M) 'S has the truth-condition that P' is not truth-conditional.

Boghossian, I surmise, would have us argue that (M) is contradictory somewhat as follows:

> (1) The predicate, '. . . has the truth-condition that P', is *robust*—since, by (M), it fails to apply to at least one type of significant declarative sentence (namely, 'S has the truth-condition that P').
> (2) The predicate, '. . . has the truth-condition that P', is *not robust*—since, by (M), predications of it do not generate sentences with substantial truth-conditions.

But this contradiction is just an artefact of description—the assumption is quite unsupported that there is any characteristic of predicates, 'robustness', which doesn't belong to a predicate unless that predicate generates truth-conditional sentences but does belong to 'true' and all cognate predicates like ' . . . has the truth-condition that P' provided they apply selectively among correctly assertible declarative sentences. The content non-factualist should merely rejoin that the contradiction lies not with (M) but with conditions imposed on 'robust' which are implicitly inconsistent with (M). All he is committed to is that predications and denials of truth-conditionality are not themselves truth-conditional.

[9] Thanks to Paul Boghossian for discussion.

References

Boghossian, P. 1989: The Rule-Following Considerations. *Mind*, 98, 507–49.
Boghossian, P. 1990: The Status of Content. *Philosophical Review*, 99, 157–84.
Churchland, P.M. 1984: *Matter and Consciousness*. Cambridge, MA.: MIT Press.
Kripke, S. 1982: *Wittgenstein on Rules and Private Language*. Cambridge, MA.: Harvard University Press.
Mackie, J. 1977: *Ethics: Inventing Right and Wrong*. Harmondsworth: Penguin.
Putnam, H. 1980: Models and Reality. *Journal of Symbolic Logic*, 45, 464–82.
Quine, W.V.O. 1960: *Word and Object*. Cambridge, MA.: MIT Press.
Shiffer, S. 1987: *Remnants of Meaning*. Cambridge, MA.: MIT Press.
Stich, S. 1983: *From Folk Psychology to Cognitive Science: The Case Against Belief*. Cambridge, MA.: MIT Press.
Wright, C. 1984: Kripke's Account of the Argument Against Private Language. *Journal of Philosophy*, 81, 759–78.
Wright, C. 1993: *Truth and Objectivity*. Cambridge, MA.: Harvard University Press.