

COMMENTS AND CRITICISMS

JOHNSTON ON HUMAN BEINGS*

MARK Johnston¹ has recently proposed a solution to the problem of personal identity, in which persons are claimed to be essentially human beings, a kind distinct from the merely biological entities that are human organisms.

After criticizing wide psychological reductionism coupled with the method of cases and proposing that we begin our theorizing with the "humble and ubiquitous practice of reidentifying each other over time" (63), Johnston gives his positive account (which I shall discuss here), which he rests on Shoemaker's case of Brown and Robinson, where Robinson receives Brown's brain in a transplant. The predominant reaction to the case, says Johnston, is that Brown is Brownson (namely, the person with Robinson's body and Brown's brain). Thus, although the alternative method will begin by reidentifying persons as human organisms, i.e., biological entities, it seems we have to depart from this approach, since, in bizarre cases such as this one, the brainless body of Brown (which Johnston calls "Brownless") is the same organism that Brown was, but Brown is now Brownson. The reason for this reaction to the case (which Johnston regards as a correct reaction) is, quite simply, a Lockean one: Brownson's tracing of himself backward in time relies on memory; memory claims by which one traces oneself essentially involve claims about personal, not merely mental, identity; Brownson's memory claims are *ex hypothesi* the same as Brown's; therefore, Brownson is Brown.

Before looking at Johnston's own characterization of persons, it must be remarked that it is not at all clear that our predominant reaction to Shoemaker's case is to suppose that Brownson is Brown and what this supposition implies, namely that Brown survives during the transplant as his disembodied brain. *If* this is our reaction and the reaction is justified, then Johnston may be right to claim that persons are not essentially human organisms; but it is quite conceivable that with extra information about brain-stem material in Brownless we do *not* have the intuition that Brownson is Brown, even though Brownson is psychologically continuous with the preoperative Brown. Our intuitive reaction to Brownless might be the same as it would be to Brown after he suffered massive brain damage in an

* I would like to thank David Pears and Derek Parfit for their helpful comments on an earlier draft of this paper.

¹ This JOURNAL, LXXXIV, 2 (February 1987):59-83.

accident, or during the terminal stages of a degenerative brain disease: of course, we might judge that the person before us is not the same person as Brown, qualitatively speaking, in which case we would, more properly, say that Brown is not the same person as he was, but we would still affirm numerical identity.² And the case raises the sorts of questions Derek Parfit discusses concerning the psychological spectrum (*ibid.*, pp. 231–233): just how much of Brown's brain must be removed before we no longer affirm that he is the organism before us? If his brain were removed in parts over a period of time, at what point would we deny numerical identity of the person? It seems that our intuitions here are by no means as clear as Johnston seems to think they are.

Now, the fact that Brownson is, *ex hypothesi*, fully psychologically continuous with Brown, cannot be relied on by Johnston to justify by itself the intuition that Brownson is Brown, since that would be to take the reductionist point, and it would then be a short step to acknowledge that *identity* did not matter. Instead, what Johnston relies on (which, together with Brownson's memory claims, yields the conclusion that Brownson is Brown) is the fact that it is Brown's own brain which has been transplanted into Robinson's head. Hence his proposal that persons are human beings which, although normally reidentified in terms of human organisms, are not the same as human organisms, since a human being can outlive the organism to which it belongs (in some sense of 'belongs'). We are then asked to believe that the person Brown, during the transplant, survives as a "radically mutilated human being," to wit a brain. Now, this two-fold characterization of human beings and human organisms is both ad hoc and obscure. Just what are these human beings? Johnston himself acknowledges his rather stipulative use of the term 'human being' (64). He certainly cannot mean what we ordinarily mean by 'human being', since what we ordinarily mean, if indeed there is any definite meaning at all, is "human organism," "member of the species *homo sapiens*." Just what Johnston's human beings are is quite unclear: a debodied brain is claimed to be a radically mutilated human being, though not a radically mutilated human organism. Presumably then, when it was embodied, the human being was not radically mutilated. Neither was the organism. So a healthy, un-mutilated member of *homo sapiens* is both a being and an organism [though the being is not a separately existing entity—Johnston wants to deny this (63, fn. 7)], and upon mutilation can be one or the other without being both. (If you take away everything but its brain, what

² See Parfit, *Reasons and Persons* (New York: Oxford, 1984), pp. 201/2.

you have left is a being without the organism; if you take away its brain, what you have left is the organism without the being.) It seems that what Johnston is getting at is that the "beingness" of the human organism is a function of its capacity to support a (complex) mental life. But we have *no* knowledge (and may never have knowledge) of the capacity of a disembodied brain to support *any* mental life: after all, *it is only when embodied that the brain demonstrates any mentality at all*. It is invalid to argue, from the fact that a human organism with a severely damaged brain has an impoverished mental life, to the fact that a brain *on its own* can support a mental life. In which case Johnston might be forced back to the position that only the whole organism, i.e., at least a body and a brain, is a person. But it is open to the reductionist to reply that, if it is the mental life of the human about which we are concerned, then it is a mere prejudice on Johnston's part to require, as a condition on Brown's survival, that some *physical* part of Brown, i.e., his brain (or, for all we know, his brain and a good deal of the rest of his body), persists through time as the bearer of Brown's mental life: it is enough if Brown's mental life itself is recreated. Why then could Brown not survive teletransportation, whereby an exact mental and physical replica of himself appeared on Mars? If mental life is what matters in personhood, as Johnston seems to think as much as the psychological reductionist, then why stop with the brain and not allow the human being (= person) to be traced, not merely through a brain transplant, but through full teletransportation? That is to say, why does the dispositional and occurrent mentality *itself* not constitute the human being, irrespective of the particular brain tissue in which it is realized?

The intuition generated by Bernard Williams's first presentation of his conundrum (whereby a doctor uses a machine to transfer *A*'s dispositional and occurrent mentality into *B*'s brain and vice versa, producing the same effects as would be produced by brain transplantation) is that *A* and *B* swap bodies when their dispositional and occurrent mentalities are swapped. This reaction is incompatible with Johnston's theory, which generates the result that *A* could only survive in what was *B*'s body if there were a brain transplantation, and so Johnston suggests that our intuitive reaction to Williams's first presentation is the result of distortion by what Johnston calls the "psychological- and social-continuer effects." Given that he is keen to begin his theorizing about personal identity from a consideration of our ordinary practice of reidentification of persons, however, it is hard to see why the psychological- and social-continuer effects are seen as *distorting* influences on the metaphysical truth of the matter. After all, such influences are a central part of our practice of rei-

identifying persons, and inform that practice. The inclination to trace *A* through a body swap and so reidentify him in the *B*-body is generated by the knowledge that the person associated with the *B*-body will be a unique psychological and social continuer of the preoperative *A*. Despite Johnston's reliance on reidentification as a "reliable and unproblematic" source of knowledge about personal identity, however, he stipulates that our practice in the case of Williams's conundrum is distorted. Why? Because it does not conform to the truth of the matter, which is that *A* and *B* have not swapped bodies. Why is this the truth of the matter? Because of the stipulation that a condition on survival is that some *physical* part of the person, viz., his brain, persist as the bearer of the person's mental life. The reductionist, however, rejects all of this. Our tendency to trace *A* through a body swap is correct, he says, not merely as a tendency in conformity with the "practical concomitants of personal identity," but as a tendency which reveals the truth that psychological continuity per se is all that matters in survival.

Johnston also claims that, in any case, there is something arbitrary in our inclination to trace *A* through a body swap: "it is difficult to get the uninitiated to come up with the body-swap intuition in response to the first presentation of Williams's case when the first presentation comes immediately after cases that emphasize the importance of bodily continuity" (81). He claims that our intuitions in such a case vary in accordance with the types of case to which the case at hand is assimilated. To explain our intuitive reaction by appealing to a survey is specious. The second presentation of Williams's case, whereby *A* is shown to survive radical psychological discontinuity (a doctor has to perform a painful operation on *A* without anaesthetic; he tells *A* he can bypass the pain by causing radical psychological discontinuity in him, and by reorganizing *A*'s dispositional and occurrent mentality after the operation so that it is continuous with *A*'s preoperative state), is said to stress the importance of bodily continuity. And yet, although presented at the same time as the first presentation, it does not seem to alter our intuition in the latter that *A* and *B* have swapped bodies. On the contrary, the whole point of the example for Johnston is that our intuitions in the two presentations stand in stark contrast!

In conclusion, apart from the apparent doctrinal difference between Johnston and supporters of reductionism (i.e., concerning teletransportation and, generally, the cut-off point for survival, which Johnston draws, unjustifiably, I claim, at the boundary of persistence and nonpersistence of a certain type of proper physical part of the person, and which Parfit draws at the boundary between

the persistence and nonpersistence of a chain of continuous mental life irrespective of the *particular* physical matter in which it is realized) it is hard to see where else any differences may lie. For it appears that, in the puzzle case of fission, where *A*'s brain is split and transplanted into the skulls of *B* and *C*, the deliverances of both Johnston's and the reductionist's theory will be the same, viz., that *A* survives as *B* and *C* though is neither one of them. What the reductionist will then point out, however, is that identity cannot be what matters in survival, for the familiar reasons. And if Johnston accepts this, as arguably he should, then his theory will not be one of personal *identity*: it will simply be a weaker form of reductionism.

DAVID S. ODERBERG

Wolfson College/Oxford

BOOK REVIEWS

Collected Papers. GARETH EVANS. New York: Oxford University Press, 1985. 407 p. \$39.95.*

Gareth Evans's *Collected Papers* contains all of his published papers. A little over half the volume deals with reference and related topics in semantics and the philosophy of language. These are the articles with which I shall be concerned.¹

In my judgment, Evans's best work in semantics is his discussion of pronouns in articles 4, 5, and 8. Two of his negative theses stand out as having lasting significance. The first is that some occurrences of pronouns with quantified antecedents, such as those in (1), do not function either as variables bound by their antecedents, or as pronouns of laziness (which are equivalent to repetitions of the antecedents).

(1a) *Few students* came to the party, but *they* had a good time.

(1b) If Mary buys a *present*, then *it* will be expensive.

* I thank Gilbert Harman, Saul Kripke, David Lewis, Stephen Neale, Lorenz Puntel, Nathan Salmon, and George Wilson for their comments.

¹ One article, "The Causal Theory of Names," is well enough known not to require extended comment. Although originally seen by Evans as a way of reconciling description theories with Saul Kripke's insights in *Naming and Necessity* (Cambridge: Harvard, 1980), it has largely been superseded by Evans's more recent *The Varieties of Reference* (New York: Oxford, 1982). In retrospect, its most salient points seem to be its "Madagascar" example of reference change and its emphasis on the role of speakers' intentions in determining and transmitting reference.