DECIDING TO BELIEVE WITHOUT SELF-DECEPTION\*

DAVID HUME noticed (as have many others) that not directly subject to our conscious control. His account suggests that it is a matter of contingent fact that one's beliefs do not respond to one's will, similar to the way in which hiccoughs do not respond to one's will. Bernard Williams' has persuasively argued that our inability to believe at will is not a contingent fact, but a consequence of the conceptual relations among the notions of belief, evidence, and truth. To believe p is to believe that p is true. But, if I could simply will myself to believe p, then I could do so whether p is true or not, and I would know that I could do so. Presumably the notion of trying to will myself to believe something would only occur to me when I find the evidence for p insufficient to warrant belief in the normal way. So, if it were possible for me knowingly to believe at will, it would have to be possible for me to believe that p is true, knowing all the while that the belief is unwarranted by the evidence, that the belief was acquired at will, and that beliefs can be so acquired irrespective of their truth or falsity. Williams rightly judges that this borders on incoherence, and concludes that "With regard to no belief could I know -or, if all this be done in full consciousness, even suspect- that I had acquired it at will" (148).

Having rightly ruled out the possibility of believing "at will, just like that," Williams grants that "there is room for application of decision to belief by more roundabout routes" (149). But the pursuit of belief by such "roundabout routes" is, he says, deeply irrational and can succeed only by means of self-deception. Presumably one would at least have to deceive oneself about the origins of one's belief - i.e., about the fact that one lacked warranting evidence for one's belief and had arrived at one's belief as a consequence of having voluntarily undertaken a program of belief inducement via roundabout routes. This is how Jon Elster2 interprets Williams's position in his recent discussion of the issue. Elster approvingly states the principle as follows: "one cannot both believe p and believe that one's belief that p stems from a decision to believe" (50). David Pears3 takes a similar position in his discussion of a hard case

\* I am grateful to Michael Phillips, Amelie 0. Rorty, Peter Skagestad, Lilly-Marlene Russow, and Michael McPherson for helpful comments at various stages of this paper's development. 1 1"Deciding to Believe," in *Problems of the Self* (New York: Cambridge, 1973).

2 *Ulysses and the Sirens* (New York: Cambridge, 1979).

3 "Freud, Sartre and Self-Deception," in Richard Wollheim, ed., *Freud: A Collection of Critical Essays* (Garden City, N.Y.: Anchor Books, 1974), pp. 97-112.

in which an individual believes not-p but finds that belief unacceptable and sets out to induce in himself a belief that p. After recounting strategies by which one might try to manipulate oneself into holding the desired belief, he states, "At some point in your progress it must become impossible for you to see through what you are doing to your motive for doing it" (108). Later he reiterates, "if this plan is going to succeed, there must come a stage in its execution at which you cannot review it in a way that brings in your motive" (ibid.).4

Each of these philosophers holds that an effort to induce in oneself a belief in the truth of some proposition that one believes to be false (or, an easier case, a proposition for which one lacks warranting evidence) can succeed only if one manages, somewhere along the way, to forget that one is engaged in such an effort. Although this view has strong intuitive appeal, I think that it is false, and that it can be shown false by example. I suspect that Williams's, Elster's, and Pears's acceptance of the view stems, first, from an underestimation of the extent to which nonevidential factors can influence belief, and, secondly, from a disregard of the fact that one's criteria for evidential warrant are themselves beliefs amenable to manipulation in "roundabout" ways. As the following example shows, it is indeed possible to believe p while believing that one's belief stems from a motivated decision to believe p, so long as one manages manipulatively to alter not only one's beliefs about the truth or falsity of p but also one's criteria for evidential warrant with regard to such beliefs.5

Nick is a bright, very ambitious biology major at a small fundamentalist college. He aspires to professional success as a biologist, and is willing to go to great lengths to attain it. He rightly suspects that his professional prospects are minimal so long as he maintains and avows his hitherto lifelong creationist beliefs. He has considered remaining a closet creationist while publicly avowing evolutionary theory, but a life of hypocrisy and deceit is an unappealing prospect. Overall, he judges that the best solution to his problem would be for him to believe that evolution is the correct theory of the development of life forms on earth. So he decides to believe it.

4 In a passage a propos of the case to be considered, Pears qualifies his claim slightly by saying: "In the later stages either my original motives must be screened, or it will come through in a version which does not include my rational tendency to believe not-p" (107). Pears does not explain this, but my example may be a case of what he has in mind. (See footnote 6.)

5 This example was suggested to me by Peter Skagestad in an unpublished paper on Nietzsche's *Uses and Abuses of History*. Overall, the ideas in this paper have developed in stimulating discussions with Skagestad.

Nick believes (and always has) that revealed scripture, taken literally, is the ultimate arbiter of disputed questions, so it will be difficult for him to induce in himself the desired belief. But he knows something about the way in which nonevidential factors influence beliefs and he works out a project of belief inducement based upon that knowledge. He knows enough about psychology to know about dissonance-reduction mechanisms of the mind which tend to bring belief into consonance with behavior. He knows enough social psychology to know the extent to which our beliefs are influenced, in nonevidential ways, by the people whom we admire and with whom we spend time. He knows enough hermeneutics and philosophy of science to know that the ways in which evidence is interpreted - and even what counts as evidence - is in great part a function of one's guiding theory and interpretive hypotheses. And, finally, he knows enough to know that our selective attention to evidence and employment of interpretive hypotheses are largely a matter of habits - habits which, like other habits, can be cultivated over time.

So, in a move designed both to further his career directly and to bring it about that he accepts evolutionary theory, he decides to go to Harvard for graduate training. Though he believes evolutionary theory to be false, he has decided, on prudential grounds, to believe that it is true, and he is confident that the prestige of his instructors, various social pressures, and the internal dissonance brought on by the need constantly to avow and utilize evolutionary assumptions (in class, in exams, etc.) will ultimately induce in him the desired belief.

According to Williams's and Elster's thesis, Nick will be unable to succeed in his project of belief inducement unless he manages to forget or to deceive himself about his motives and about the origins of his belief. On this view, so long as he remembers that his belief is a consequence of his having employed a roundabout strategy of belief inducement, relying on non-evidential mechanisms and influences, he will be unable really to believe in the truth of the theory of evolution. But to see the falsity of this thesis, we need only fill in the rest of the story:

Six years later, Nick emerges from graduate school, Doctorate in hand, fully versed in and accepting of contemporary biological theory. He remembers quite well the entire story of how he had decided to come to Harvard in hopes that he would be seduced into believing that which he then thought to be false. He smiles inwardly at the benighted bumpkin he was - so benighted that he had thought evidential considerations insufficient to warrant belief in the truth of something as obviously true as evolutionary theory.

The point of the story is that Nick has successfully carried out his decision to believe p and is fully aware that his belief is a consequence of a strategically planned "roundabout" program of belief inducement. He has not forgotten anything, and has no need to deceive himself about anything. Thus, Williams's and Elster's thesis is shown to be false.

One might argue that Nick has indeed forgotten something - he has forgotten that revealed scripture is the final arbiter of disputed questions. But, in fact, he has

not forgotten this; he holds simply that it is false and that he was mistaken when he held it to be true. Perhaps, then, he has forgotten the evidence that once led him to hold his prior beliefs. But, in fact, he has not forgotten it at all; it is just that now he is inclined to think that evidence very weak and to think that his prior beliefs were mostly without evidential support, resting almost entirely upon ignorance, social pressure, superstition, and unjustified credence given to the views of his fundamentalist parents and teachers. In short, he has not forgotten anything about the past, though his present account of his prior epistemic state is different from the account that he would have offered previously. The story that he now tells about his previous epistemic position is a redescription of that position, but not one which requires either selective forgetfulness or self-deception.

It is important to note that there is a kind of symmetry between Nick's earlier perspective upon his later position and his later perspective upon his earlier position. Referring to him prior to implementation of his strategy of belief inducement as Nick1, and upon graduation from Harvard as Nick2, this symmetry can be described as follows: Nick1 views his own beliefs as based upon warranting evidence and foresees Nick2's beliefs to be the results of nonevidential causal influences (professors' prestige, social pressures), prejudicially selective habits of attention, and employment of tendentious interpretive hypotheses. Likewise, Nick2 views his own beliefs as evidentially warranted and those of Nick1 as the result of ignorance, distorted perspective, ritual belief reinforcement and other nonrational causal influences. The fact that each can "explain away" the beliefs and the evidential standards of the other provides sufficient "epistemic insulation" between the two to permit Nick2 sincerely to believe p without having to forget or deceive himself about the fact that his belief is a consequence of Nickl's decision to believe p. Perhaps this is what Pears had in mind when he suggested that "my original motive must be screened, or . . . come through in a version that does not include my rational tendency to believe not-p" (107).6 In Nick's case, his original motive "comes through in a version" that includes a memory of having had a tendency to believe not-p and having thought it rational, but his present view of that tendency holds it not to have been rational. (Or, if perhaps it was rational given his evidence at the time, that evidence was woefully inadequate - so inadequate that it led him to distorted notions of rational warrant.)

6 See above footnote 4.

In a sense, the suggested counterexample to Williams's thesis is but an application of the Duhem-Quine thesis, according to which one can hold any proposition true if one is willing (and able!) to make sufficient changes elsewhere in one's system of beliefs. Williams might hold that this sort of case is devoid of practical interest precisely because the individual, having had to change his standards of evidential warrant, has also had to change innumerable other beliefs -and this, it might be thought, would be so disruptive that no one would choose to do it. In response, it should first be said that this objection does not affect the conceptual import of the case. Moreover, the extent to which such strategic belief inducement will require alteration of other beliefs will vary widely from case to case. In some cases, massive overhauling of one's belief system might be required; in other cases, the effects of the change might be predictably slight or easily isolable.7

I would suggest, in conclusion, that rather than think of what Nick has done as a case of self-deception, we would do better to construe it along the lines of conscious character planning on an Aristotelian model. Suppose that I do not enjoy some x, but think that I would be a better person if I did. So I set about to engage in certain patterns of action, the predictable result of which will be my coming to enjoy x. In time, a *hexis* develops and I find myself enjoying x. In M. F. Burnyeat's8 felicitous phraseology, I learn to enjoy x where that is "not sharply distinct from my learning that x is enjoyable." A cynic might say that I had deceived myself about the enjoyability of x, but a more charitable account would hold that I had successfully carried out a project of conscious character planning.

Nick's motives are perhaps not so noble, but the structure of the case is comparable. He set out to engage in a pattern of action, the predictable result of which was his coming to believe a certain theory in biology. In time, a *hexis* developed, and he found himself believing the theory. He learned to believe the theory, where that is not sharply distinct from learning that the theory is believable.

7 It is, after all, quite possible to maintain different criteria of evidential warrant with regard to different aspects of one's life or different subject matters. All that is needed is either failure to notice that one is employing different criteria or some (perhaps quite vague) cover story about the appropriateness of the disparate criteria. Consider also recent suggestions that loyalty consists in part of a willingness to apply more stringent criteria of evidential warrant with regard to unflattering or maligning claims about those to whom one is loyal.

8 "Aristotle on Learning to be Good," in A. 0. Rorty, ed., Essays on Aristotle's Ethics (Berkeley, Calif.: University Press, 1980), p. 76.

One might judge that Nick has deceived himself, but a more charitable account would hold that he has successfully carried out a program of belief acquisition via a roundabout route.9

Most of us would probably agree that Nick moved from an unwarranted belief in a false view to a warranted belief in a true one. But our judgment of the force of the example should not be thought to depend on this fact. If Nick had started out as an evolutionary biologist and had had sufficiently strong prudential motives for wanting to believe some fundamentalist/creationist doctrine, he could have done so, and could have done so without self-deception. One's sense for what is believable (otherwise known as one's criterion for epistemic warrant) can be altered in any number of nonrational ways. To the extent that one manages to bring about such an alteration, one's beliefs can be consciously manipulated without the need for self-deception.

J. THOMAS COOK Rollins College

9 As an aside, I might mention that those with Parfitean inclinations could interpret the apparent absence of self-deception in this case in a different way. It might be thought that there is no single Nick who is both deceiver and deceived, since Nickl's machinations not only produced certain beliefs, but brought into being a different (later) self to hold those beliefs. Thus, though Nick, did willfully bring it about that Nick2 hold beliefs that Nick1 held to be false, the deception involved here is not self-deception, for the deceiver and the deceived are quite distinct selves. I have reservations about this way of talking, but the epistemic insulation between Nick1 and Nick2, coupled with the rather patronizing story that each has to tell about the other's epistemic status, suggests that they are ripe candidates for being considered different selves.