

An Empirical Basis for Psychological Egoism

Author(s): Michael Anthony Slote

Source: *The Journal of Philosophy*, Vol. 61, No. 18 (Oct. 1, 1964), pp. 530-537

Published by: Journal of Philosophy, Inc.

Stable URL: <https://www.jstor.org/stable/2023495>

Accessed: 20-12-2018 21:35 UTC

JSTOR is a not-for-profit service that helps scholars, researchers, and students discover, use, and build upon a wide range of content in a trusted digital archive. We use information technology and tools to increase productivity and facilitate new forms of scholarship. For more information about JSTOR, please contact support@jstor.org.

Your use of the JSTOR archive indicates your acceptance of the Terms & Conditions of Use, available at <https://about.jstor.org/terms>



JSTOR

Journal of Philosophy, Inc. is collaborating with JSTOR to digitize, preserve and extend access to *The Journal of Philosophy*

disclose is true) and discipline (meaning diligent exercise of the will in the attempt to work out the implications of the disclosures for the living of life in the everyday, common-sense world).

Nowhere today in Western civilization are these two conditions jointly fulfilled. Churches lack faith in the sense just mentioned; hipsters lack discipline. This might lead us to forget about the drugs, were it not for one fact: the distinctive religious emotion and the emotion that drugs unquestionably can occasion—Otto's *mysterium tremendum, majestas, mysterium fascinans*; in a phrase, the phenomenon of religious awe—seems to be declining sharply. As Paul Tillich said in an address to the Hillel Society at Harvard several years ago:

The question our century puts before us [is]: Is it possible to regain the lost dimension, the encounter with the Holy, the dimension which cuts through the world of subjectivity and objectivity and goes down to that which is not world but is the mystery of the Ground of Being?

Tillich may be right; this may be the religious question of our century. For if (as we have insisted) religion cannot be equated with religious experiences, neither can it long survive their absence.

HUSTON SMITH

MASSACHUSETTS INSTITUTE OF TECHNOLOGY

AN EMPIRICAL BASIS FOR PSYCHOLOGICAL EGOISM

IT is commonly believed in the philosophical world today that the age-old problem of psychological egoism is merely a pseudo-problem and that this is true just because the a priori philosophical arguments that have traditionally been given in favor of egoism depend in the main upon confusions about the logic of our ordinary language. It has been claimed, for example, that the well-known argument that we act selfishly even when we want to help others because in such cases we are still attempting to *satisfy our own desire* to help others, is fallaciously generated by misunderstandings of the proper use of terms like 'want', 'satisfy', and 'desire'.¹

In *Butler's Moral Philosophy*, Austin Duncan-Jones, expressing Butler's view, and, it seems from the context, his own as well, states that if there is something wrong with all the a priori philosophical arguments that have traditionally been given in favor of egoism (which he has earlier identified with the doctrine that

¹ See P. H. Nowell-Smith's *Ethics* (New York: Philosophical Library, 1957), ch. 10, *passim*.

all human acts are selfish),² then there is little else to recommend the theory, since "the appearance of things, undistorted by theory," is that men sometimes do act unselfishly, disinterestedly. Only one who already believed in the validity of the philosophical arguments for egoism would have any reason to interpret the facts of human behavior in a way compatible with the doctrine of egoism.³ Thus Duncan-Jones seems clearly to be ruling out the possibility that the (empirical) facts as they stand could, with any semblance of objectivity, be used to support egoism. And many other contemporary philosophers would, I think, tend to agree with him.

In the present paper I wish to argue that psychological egoism may well have a basis in the empirical facts of human psychology. Certain contemporary learning theorists, e.g., Hull and Skinner, have put forward behavioristic theories of the origin and functioning of human motives which posit a certain number of basically "selfish," unlearned primary drives or motives (like hunger, thirst, sleep, elimination, and sex), explain all other, higher-order, drives or motives as derived genetically from the primary ones via certain "laws of reinforcement," and, further, deny the "functional autonomy" of those higher-order drives or motives.⁴ Now it is a hotly debated issue in contemporary Learning Theory whether any theory such as we have described briefly above could adequately explain adult human behavior. I shall, however, argue only that a theory of the above kind may well be true, and that from such a theory, fortified only by one additional psychological premise, the truth of egoism (non-altruism) logically follows. I hope to show, thereby, that the question of psychological egoism is still an open empirical issue, however fallacious be the *philosophical* arguments for it.

But what is "functional autonomy," and how does the lack of it help to show our actions to be selfish? According to behavior-

² It has been suggested to me by P. R. Foot that only those of one's acts which are somehow related to the wants or interests of others can correctly be called either selfish or unselfish. If this be so, then Duncan-Jones' definition of egoism will make that doctrine trivially false, just because there are some human actions that are neither selfish nor unselfish. In order to avoid such an eventuality, I shall mean by *egoism* the slightly different thesis (perhaps more accurately, but clumsily, designated *non-altruism*) that no human act is ever *unselfish*.

³ See p. 109 of *Butler's Moral Philosophy*.

⁴ It will not, I think, be necessary for my purposes to be truer to ordinary language or more precise with the concepts of drive and motive than are the learning theorists themselves. Thus, e.g., I shall be using 'drive' and 'motive' interchangeably in this paper.

istic learning psychologists a higher-order (acquired) motive is functionally autonomous when it becomes causally independent of primary motives (especially of those motives association with which enabled it to be acquired in the first place) in such a way that one will indefinitely keep acting from that motive, even if rewards for those other, primary, motives are no longer in general associated with such action.⁵ A higher-order drive or motive is *not* functionally autonomous, i.e., is functionally dependent, if and only if when we cut off all reinforcement of it by primary rewards (rewards of primary drives) and there are, in addition, both a sufficient number of "extinction trials" (occurrences of acts done from that higher-order motive which are not associated even indirectly, i.e., through other higher-order motives, with primary rewards) and a complete absence, during those extinction trials, of primary rewards for any similar higher-order motives (to eliminate the possibility of generalization of primary rewards from motives other than that being extinguished), the higher-order drive or motive actually does extinguish; i.e., the person whose higher-order motive is being extinguished eventually, even if perhaps only very gradually, ceases to act from that higher-order motive.

It is necessary for a motive *derived genetically* from "selfish" (or at least not *unselfish*) primary drives also to be *functionally dependent* upon them if we are to be able to say that acts performed from that motive are never unselfish. For the fact that in the past we performed such acts only because they led to the satisfaction of some other non-unselfish motive or motives, i.e., because they were reinforced by primary-drive rewards, does not show that such acts performed *now* are not unselfish. To argue thus would be indeed to commit a "genetic fallacy." An act must *presently* be causally connected with drives that are not unselfish in order to be considered selfish. Now those who deny functional autonomy are saying in effect that whenever, e.g., one acts benevolently (i.e., from what the psychologist would call the higher-order motive of benevolence), one would not be performing that act, or, at least, would not indefinitely continue to perform acts of that kind, if such benevolent action were not in general still associated with and reinforced by the satisfaction of such non-unselfish primary drives as hunger and thirst, whether those drives be the

⁵ I have given a very brief account of the notions of primary and higher-orders drives or motives; but it should be sufficient for the purposes of this paper. A more complete account of these and of the other psychological notions I make use of can be found in practically any textbook of experimental psychology.

same as or different from the primary drives from which the motive of benevolence actually originated.

I do not, however, wish to maintain that the hypothesis of functional dependence (together with its learning-theoretical underpinnings) entails egoism all by itself. The hypothesis does, indeed, entail that we would never continue to act "benevolently" or "self-sacrificingly" if such action on our part were not in general reinforced by the rewarding of selfish primary drives. But is it not possible that the primary rewards received, in general, when one acts benevolently or self-sacrificingly are not so great as those relinquished in the doing of such acts? It might, in other words, be the case that a poverty-stricken mother who sacrificed some of her own food so that her child might eat better would not act in this way unless she were, in general, receiving (however indirectly) some primary-drive satisfactions for her sacrifices. And yet we would still call her actions unselfish if we thought that the rewards she was sacrificing (reduction of her hunger) were greater than those she was getting in return; for is not the habit of giving more than one asks in return an exemplary case of unselfishness?

I should like now to show that a certain empirical hypothesis may, when taken together with the hypothesis of functional dependence, entail the thesis of psychological egoism and rule out the possibility of a case like the above, even if the hypothesis of functional dependence taken alone does not.⁶

Let us imagine that we have a method for determining empirically which primary rewards a person prefers to which others. We set up various situations where the man has to choose between primary rewards, situations involving no moral factors and no interests of other people, and determine the man's preferences. A learning theorist might claim that it is true as a matter of empirical fact that whenever a man systematically (i.e., as a general rule) continues to sacrifice primary reward x to other people, he does so only because he usually obtains thereby some primary reward y and because y ranks higher than x on the person's preference scorecard, as determined in situations where no considerations of other people's interests and thus of sacrifice to other people's interests were involved.⁷ And the above empirical claim, which involves, but is not exhausted by, the claim that functional de-

⁶ I am again indebted to Mrs. Foot for the insight that the hypothesis of functional dependence does not itself, alone, entail that no act is unselfish. In addition, I am indebted to discussion with Prof. R. P. Wolff for some of the points I shall be making hereafter.

⁷ Of course, there are some primary rewards, like sexual gratification, that are very hard to measure in isolation from all moral considerations.

pendence is true, entails, I think, the thesis of psychological egoism. For if our conscious acts of benevolence and sympathy and sacrifice, etc., would eventually cease to be performed by us if we did not, in performing those acts, in general give away *less* in the way of primary-drive satisfactions than we actually got in return, the inevitable conclusion would be that all our acts were fundamentally (or ultimately or "really") motivated by our "selfish" primary drives. And it would be accurate, if this were the case, to say that the driving forces behind all our so-called higher actions were "selfish" primary motives, since only those higher actions would continue to be performed which usually led to a net gain of primary-reward satisfaction for the individuals performing them. If this were the case, then, indeed, none of our actions would ever "really" or fundamentally or ultimately be unselfish. The above-described case of the mother who gives away more in the way of primary-drive satisfaction than she gets for herself in return would just never come up. Psychological egoism would be true.

We have thus shown that egoism (in our sense) would be true if certain psychological hypotheses turned out to be true and that the question of the truth of psychological egoism is an empirical question. But that is not to say that contemporary psychology has been able to prove the truth of these hypotheses or that psychologists are even all agreed that, with the further advancement of psychology as a science, these hypotheses will as a matter of fact be verified. There are many psychologists who think, for example, that some higher-order drives *do* become functionally autonomous. Gordon Allport, for instance, has brought to light a good deal of psychological evidence in favor of this contention.⁸ Furthermore, the hypothesis of functional dependence is very difficult to establish experimentally, for reasons well known to psychologists. In the words of Neal Miller, "a strong learned drive may seem unaffected for many [extinction] trials and still eventually extinguish. When generalization, higher-order reinforcement, and shifts from one reinforcing agent to another are added to this possibility, it can be seen how difficult it is in complex human situations to determine whether a habit [drive] actually is functionally autonomous."⁹ In other words, even if there is no functional autonomy, there are many ways in which

⁸ See his *Personality: a Psychological Interpretation*.

⁹ "Learnable Drives and Rewards," in S. S. Stevens, ed., *Handbook of Experimental Psychology*, p. 469.

See also D. C. McClelland, "Functional Autonomy of Motives as an Extinction Phenomenon," *Psychological Review*, 49 (1942): 272-283.

a functionally dependent drive might *appear* to be autonomous, because of distorting psychological factors that can never with absolute certainty be ruled out in the context of human motivation. However, the question whether some drive is autonomous is still empirical in principle, however difficult it may be in practice, given the current rudimentary state of the science of psychology, to determine whether that drive would extinguish if its association with primary reinforcements were entirely severed.

Consider also the hypothesis that, together with the hypothesis of functional dependence, entails egoism, the hypothesis, namely, that even if people sometimes do sacrifice and continue to sacrifice a certain kind of primary reward, they never do so unless they in general get some greater primary reward in return. This hypothesis, I think, might also turn out to be empirically false. It might turn out, for example, that, although people never made sacrifices unless they got something in return, they sometimes sacrificed some primary reward x for some other primary reward y even though y ranked lower on their preference scorecard as determined in morally neutral circumstances. In such a case there are two things we can say. We might well say that the scorecard as determined in morally neutral circumstances does not tell us a man's real preferences, for if he prefers x to y in neutral circumstances, but prefers y to x when certain other people's interests are involved, who can say what his real preference is? We might, on the other hand, want to say that what a man prefers in morally neutral circumstances really does tell us fairly accurately what he *really* prefers, so that if his preferences differ where moral considerations are involved, we have a right to say that the man has, in the interests of morality, gone against his own preferences and made an unselfish sacrifice. Whether we should want to say the first or the second of these things would depend a great deal, I think, on a number of other scientifically relevant factors. Anyone who would *in principle* refuse to say the second kind of thing would in effect be considering it to be tautologously true that men do not persist in acting against their own self-interest; that is, he would be making the thesis of egoism into a mere tautology, which, I prefer to think, it is not. It seems that there very well could be circumstances in which it would from a scientific point of view be advisable to say that a man had acted against his own real preferences, had persistently sacrificed a greater for a smaller primary reward out of a sense of duty or a feeling of benevolence. Such circumstances might exist, for example, if we had a detailed knowledge of brain physiology which showed that the brain contained a "preference" center and a "morality" center and that the moral-

ity center affected our actions not by directly influencing the preference center, but, rather, by acting as an inhibitor or as a modifier on certain motor impulses sent out by the preference center. Such a physiological theory would make it eminently plausible, I think, to believe that what we did in moral contexts might consistently go against our real preferences. In terms of such a theory, then, it might be possible empirically to refute the hypothesis that we never consistently or systematically sacrifice the greater for the smaller primary reward. Thus it would seem that both parts of that psychological theory which, I have claimed, entails psychological egoism, are open to empirical refutation, as well as confirmation.¹⁰

The psychological theory I have been describing should not be confused with a certain theory of human behavior and motivation put forward in recent years by the psychologist A. H. Maslow, a theory which, I believe, does *not* entail psychological egoism. According to Maslow people will not act from such higher motives as benevolence and love unless certain lower needs like hunger, safety, and elimination have already to some degree been satisfied. But once physiological and other lower needs are satisfied to a reasonable degree, needs to be benevolent, creative, loving, self-sacrificial, and the like will spring up of their own accord. And one will continue to act benevolently, creatively, etc., just as long as one's lower needs remain satisfied, even if none of one's benevolent or creative activities is actually reinforced by the satisfaction of lower needs (primary drives).¹¹

Clearly this theory differs substantially from the one I have been describing. For it does *not* assert that we will persist in acting benevolently, etc., only if such acts are in general associated with the satisfaction of selfish primary drives. It says merely that we require a certain amount of primary-drive contentment if we are to become people who constantly act benevolently *whether we are rewarded for doing so or not*. According to Maslow, a man often will habitually act from benevolence even though there is "nothing in it for him." It is clear, then, that his theory does *not* exclude the possibility of unselfish human action. The theory *we* have been discussing, on the other hand, does exclude that possibility, just because it implies that we never persist in performing any kind of action unless there is in general something in it for us.

¹⁰ There is indeed still another way we have not yet mentioned in which this psychological theory might empirically be refuted, namely, if some theory (like Hume's) which made benevolence or other unselfish motives into basic human instincts (primary drives) turned out to be correct.

¹¹ See Maslow's *Motivation and Personality*, ch. 5, *passim*.

It would seem, then, that, as psychology stands today, there is at least some reason to think that the psychological theory we have been discussing may be true. Consequently, the truth of psychological egoism is still an open empirical question. Duncan-Jones and others are mistaken in their belief that, now that the a priori arguments for egoism seem to have been shown to be fallacious, no further case can possibly be made in its favor. Perhaps the only reason philosophers are thus mistaken is their ignorance of contemporary Learning Theory, its issues, and its results.¹² It is interesting to note, furthermore, that it is impossible to object to an empirical argument for egoism, the way one so often objects to a priori arguments for egoism, by saying that such arguments end up depriving 'selfish' of the logical possibility of a contrast, thus rendering the word meaningless. For egoism will be false if either part of the psychological theory we have been discussing is false. Thus, in making the truth of egoism depend on the truth of an empirically falsifiable psychological theory, I am leaving open at the very least the *logical* possibility that egoism is false, that some acts are unselfish.

I might add, finally, that the explanation I have attempted to give of the possibility of arguing on an empirical basis for egoism may help us to understand why so many people, especially beginning students of philosophy, are so dissatisfied by attempts to discredit egoism by showing the invalidity of the traditional philosophical arguments that have been put forward to prove it, and why the doctrine of egoism keeps cropping up, however many be the philosophical voices that seek to silence it. I am willing to conjecture that egoism will not lie dead, because people in some way see that there may be more in favor of egoism than a *priori* arguments. It is my very tentative suggestion that the reason for this may be that even those with little or no training in psychology believe, however inarticulately, that something like the psychological theory we have been discussing in this paper may well be true, believe that men who act consistently in a benevolent manner, for example, would not be acting benevolently unless their selfish desires and/or interests were usually satisfied by their doing so.

MICHAEL ANTHONY SLOTE

HARVARD UNIVERSITY

¹²I do not, however, wish to suggest that *psychologists* have been totally unaware of the philosophical consequences of their own theories. See, for example, Allport, *op. cit.*, p. 206.