

cussed at all. Thus I am caught in an apparent paradox: if what I say is true, then by my own arguments I ought not to believe it. It does not make sense to be the only conformist, as it were, in town. But there is really no paradox here, since I desire to assert my thesis only hypothetically—not as a report of my own firm beliefs, but as a mere, suggested avenue of inquiry for others to pursue and evaluate. I have first-order reasons to believe in my conclusions, which I have given above. If I did not find the view pretty convincing myself, I would not have bothered writing it all down. But I really do not know what other philosophers are going to think about these arguments in the long run. If there are glaring errors, I may find that out pretty soon. But if my thesis is true, I doubt that I shall ever know this, because most other philosophers will not agree very quickly, and I already know that my own opinion in philosophy, while modestly reliable as such things go, is not *especially* to be trusted.

Perhaps the only point on which there is a true consensus among philosophers is the principle of epistemic self-reliance itself—which is what I have been questioning. But even if I were attacking this principle head-on (for example, by arguing for a straight, majoritarian conformism), I would still be obeying it, since this attack would be my own idea.²⁷ In any case, it has not been my purpose to attack the idea of intellectual autonomy itself. Instead, I have been trying to say what it really amounts to, and how it is best to be employed—that is, in the experimental, “brainstorming” spirit of contemporary physics. So my position is a little awkward, perhaps, but no more paradoxical than that of any other intellectual who is working with the proper attitude toward his own, tentative results.

THEODORE J. EVERETT

State University of New York/Geneseo

²⁷ It is more common to encounter the implicit conformist, who unquestioningly asserts the principle of “thinking for yourself,” just because everybody else does. The *locus classicus* for this point is the famous balcony scene in Monty Python’s film, *The Life of Brian*, where the reluctant savior Brian commands his followers to think for themselves, and they shout their agreement in unison.

BOOK REVIEWS

Unto Others: The Evolution and Psychology of Unselfish Behavior. ELLIOTT SOBER and DAVID SLOAN WILSON. Cambridge: Harvard University Press, 1998. 394 p. Cloth \$33.50, paper \$17.95.

Unto Others is an intellectual joyride, by turns speculative and pioneering, cautious and carefully argued, scholarly and judicious, self-righteous and strident, generous and evenhanded, and defensive and partisan in advancing its central themes: the undeserved neglect of group selection as an important mechanism of evolutionary change making possible the evolution of altruistic traits (chapters 1-3), the special significance of this mechanism for the evolutionary history of human beings (chapters 4-5), and the evolutionary considerations favoring the inclusion of altruistic impulses among our ultimate psychological motivations (chapters 6-10).

Elliott Sober and David Sloan Wilson have been putting group selection’s best foot forward for quite some time now, and they effectively confront a number of confusions about the issue, most importantly, what they call the *averaging fallacy*: defining altruism and group selection out of existence by averaging the fitnesses of genes, traits, organisms, and the like across the various environmental circumstances they encounter (including membership in groups with varying proportions of altruistic members), noting that it is the entities with the highest *average* fitness which evolve, and ignoring the level of biological organization at which the relevant selective processes actually take place. Using now-familiar examples (disease virulence, female-biased sex ratios, meiotic drive, and so on), they show how selective pressures can coexist and even conflict at different levels of biological organization: most important for them, even an altruistic trait conferring a selective disadvantage upon its bearers within each isolated subgroup of a population will increase in frequency in the global population through a competing process of selection at the level of groups themselves by the operation of Simpson’s paradox—here that groups containing higher proportions of altruists contribute more offspring to the next generation of the total population even as the frequency of altruists declines within each group—if the group benefit is sufficiently large relative to the individual cost, *if* there is sufficient variation in the proportions of altruists present in the various groups, and (most awkwardly) *if* the population is structured into isolated groups that mix and reassert

(again with sufficient variation in the proportion of altruists within each group) or otherwise compete in the formation of new groups. It is disappointing that Sober and Wilson do not provide a wealth of new empirical examples in which these conditions are clearly realized, however (especially since they suggest that such examples abound once our conceptual blinders are removed), because even their detailed and elegant discussions of these familiar examples (and convincing criticisms of earlier analyses of the conditions required for group selection by Sewall Wright and John Maynard Smith) do little to rebut the standard assessment of their position among evolutionary biologists: that defenders have shown selection at the level of groups to be a theoretical possibility, but one requiring such unusual conditions as to be rare and comparatively unimportant in nature. Of course, part of their agenda is to advertise even this much progress for group selection to scholars outside evolutionary biology.

Furthermore, Sober and Wilson do offer an implicit reply to this appraisal, arguing that accounts of the evolution of altruistic traits traditionally considered widely applicable alternatives to group selection—inclusive fitness, evolutionary game theory, and selfish gene theory—simply offer different *perspectives* on the *very same process* of group selection described by their own multilevel selection account. This conclusion depends, however, on their controversial conception of a “group” as absolutely any set of individuals whose interactions affect each other’s fitnesses, a proposal which has led many reviewers to suggest (mistakenly, I think) that the only issues here are semantic and that the authors try to win an argument over group selection simply by redefining it. While Sober and Wilson are right to insist that evolutionarily significant groups are defined by the pattern of interactions among individuals they make possible (rather than spatial or temporal boundaries), their bid for generality nonetheless rests upon supposing that this liberal conception of group selection will be a helpful or perspicuous way to think about the causes of evolutionary change in an implausibly broad range of cases. They invite us, for example, to conceive of a population randomly pairing up for rounds of prisoner’s dilemma as a single global population that is repeatedly divided into many ephemeral groups of size two and in which group selection runs rampant: while cooperative behavior is exploited and therefore selectively disadvantageous within mixed groups (pairs) of cooperators and defectors, groups with higher proportions of cooperators (100% versus 50% versus 0%) ultimately contribute more offspring to the global population (note that this is true of groups containing 50% altruists (that is, one altruist) only because of the greater productivity of the *selfish* member). Although the numbers

can be correctly crunched under this description, our understanding of why some strategies evolve and others do not is surely better advanced by the game theorist’s usual account, that is, by noting the payoff to each strategy against each potential interactor and the rate at which each is encountered—in a word, by calculating the *average* payoff each strategy receives when played against the population at large. The game-theoretic description’s explanatory advantages become still more evident when we compare the two accounts’ competing characterizations of the further condition required in order for the cooperative strategy to evolve here: for the game theorist, cooperators must be more likely to interact with one another (whether because of population structure or viscosity, the association of kin, communication, the recognition of like-minded souls, or whatever), while Sober and Wilson must represent this same condition as an increase of variation in the composition of their ephemeral “groups.”

Thus, even if we accept Sober and Wilson’s suggestion of theoretical equivalence between multilevel selection theory and its apparent competitors (and are consequently reluctant to deny *categorically* that these cases could be regarded as instances of something plausibly called “group selection,” even if it is not quite the classical group selection attacked by Williams, Maynard Smith, and others) the central issue becomes how often it will be *useful* to think of this single process as a form of group selection. Despite repeated assurances, Sober and Wilson provide few concrete reasons to think that multilevel selection theory will provide a broadly perspicuous or helpful perspective, even (as the example of two-player prisoner’s dilemma with correlated interaction illustrates) in all those cases in which population structure plays an important evolutionary role.

Fortunately, these considerations do not undermine the interest of the remainder of the book, for *Unto Others* goes on to place human beings centerstage, and Sober and Wilson go on to argue that ancestral populations of human beings did satisfy the special requirements of traditional group selection, amplified by the fact that our cognitive sophistication enabled altruists to selectively interact more easily, made it possible to evolve mildly altruistic systems of rewards and sanctions that enforced more dramatically altruistic behaviors, and rendered competing human groups increasingly distinct and internally homogeneous through the influence of cultural selection. Sober and Wilson’s anthropological evidence (drawn from a random sample of anthropologists’ field descriptions of human cultures in the Human Relations Area File) ground arguments that are among the most speculative in the book, but they are also some of the most interesting, and they make at least a *prima facie* case deserving

serious consideration (nor do the authors claim more) for the importance of group selection and the ubiquity of altruism in human evolutionary history.

The second half of *Unto Others* turns to questions of psychological (as opposed to evolutionary) altruism, contrasting the case for psychological egoism (the claim that our ultimate psychological motives are uniformly self-directed) with a pluralism that includes ultimate (rather than merely instrumental) motives that are genuinely altruistic (that is, directed toward promoting the welfare of others). They effectively demolish weak philosophical and empirical arguments on each side (including misguided appeals to parsimony) and seek to break what they regard as the ensuing stalemate with evolutionary considerations: using parental care as their example, they argue that considerations of reliability (in advancing fitness), availability (in ancestral conditions), and efficiency (in energetic cost) jointly favor the evolution of a motivational pluralism, including genuinely altruistic motivations over the egoistic alternatives.

Sober and Wilson's evolutionary argument errs, however, in treating hedonism as the sole representative of psychological egoism. This is defended by pointing out that versions of egoism which appeal to desires for external goods invariably wind up invoking hedonistic motives to explain away apparently altruistic actions, but this simply does not support the authors' conclusion that an *equally* exclusive commitment to hedonistic motives must be the most difficult form of egoism to refute! Indeed, the toughest opponent for the thesis of genuine psychological altruism, especially in an evolutionary context, is a *pluralistic* egoism in which our self-oriented impulses include a wide variety of internal *and* external cues that are (or were in ancestral conditions) relevantly connected to our fitness. This way of thinking about human motivation is familiar from the literature of sociobiology and evolutionary psychology, and it is especially important for the authors to consider, given that the core of their argument for motivational pluralism is that a genuinely altruistic psychological mechanism for ensuring parental care will enjoy greater reliability *in promoting fitness* than the hedonistic alternative they consider.

Setting aside their choice of target, Sober and Wilson do not take seriously enough the question of whether genuinely altruistic psychological motives would be easier or more likely to evolve than hedonistic counterparts. After all, parental care is neither the only nor even the most important behavior that well-adapted organisms need to perform, so any organism that engages in it will need a mechanism for *trading off* its impulses toward parental care with those directed toward other important behaviors. While Sober and Wilson argue (in

connection with comparative efficiency) that even a hedonist requires some mechanism for trading off the various *kinds* of pains and pleasures that make up her motivational structure (including the propositionally-mediated variety hedonists will need even to treat other-directed desires instrumentally), it remains reasonable to think it more likely for an organism to evolve motivations for an important new behavior that fits into the (comparatively) common currency of an existing hedonistic mechanism for trading off each of its competing motivations with one another (the most likely ancestral condition, as Sober and Wilson recognize at least for the sake of argument) than to evolve an entirely new mechanism for trading off a new fundamental motivation against each of them.

Ultimately, this pattern is representative of *Unto Others* as a whole: Sober and Wilson offer a trail-blazing blend of insight, speculation, evidence, argument, and conjecture which will surely fuel serious thought and discussion in a wide variety of fields for a long time to come, even as the raw ambition of their case repeatedly leaves its supporting argumentation lagging behind, gasping and racing to catch up.

P. KYLE STANFORD

University of California/Irvine

The Bounds of Agency: An Essay in Revisionary Metaphysics. CAROL ROVANE. Princeton: Princeton University Press, 1998. ix + 260 p. Cloth \$45.00.

The aim of Carol Rovane's *The Bounds of Agency* is to substitute an adequate and coherent conception of personhood for the confused mixture of animalism and Lockean psychologism which our intuitions encompass. The assumption governing this effort is that despite the lack of agreement over what follows if an entity is discovered, or deemed to be a person, the possession or acquisition of the property of personhood is valuable. Although it cannot be better for objects in general to be persons rather than things, it is better for human beings to be persons than nonpersons. Criteria for personhood have to support the intuition of superior value; at the same time, it is risky to load the notion with too many duties or privileges held to be knowable a priori and necessary. Rovane occupies the middle ground by claiming that "person" is a normative notion, though it is neutral with respect to specific theories of obligation. Unlike P. F. Strawson's "descriptive" metaphysics of persons which reassured readers that