Review Essays & Reviews

Reasons and Persons. By Derek Parfit. Oxford: Clarendon Press. 1984.

Traversing *Reasons and Persons* is the philosophical analog of a drive along the Southern California freeway system. Traffic in ideas is bumper-to-bumper, routes merge into each other with little prior notification, the interchanges leaving one breathless, and the novice motorist might do better to practice on more forgiving roadways. On the other hand, the vast proliferation of paths really does hang together, and one genuinely is led to marvel at the fact that the mind of man succeeded in producing so stunning a design.

  This is a *big* book. Its 543 pages include four major sections, each of which could constitute an independent volume, 10 appendices, 27 pages of endnotes, a bibliography, and an index of names. (There is, however, no subject index, an unfortunate omission in a book so rich in interconnected ideas.) Part One discusses ways in which theories can be self-defeating, with special attention to S, the Self-Interest theory of rationality; C, Consequentialism of a broadly utilitarian sort; and M, Parfit's rendering of Common-Sense morality. Part Two, "Rationality and Time" places S under the diagnostic lamp once again, this time specifically targeting the claim that rational self-interest entails equal concern for all temporal stages of one's life. Part Three concerns personal identity, an area of inquiry Parfit has notably advanced in his earlier work.[1] He argues that personal identity is nothing over and above nonbranching psychological connectedness and continuity, and that, contrary to standard belief, whether some future individual will be *me* (rather than someone who merely resembles me closely along salient psychological dimensions) does not very much matter. Finally, in Part Four Parfit addresses the question of which moral principles should guide our decisions when what we do will not only affect the welfare of future individuals but will causally determine which and how many persons will come to be.

  Such a book is both the bane and bonanza of reviewers. It is impossible even to mention, let alone examine in depth, all the important issues on which Parfit shines the spotlight of his formidable analytical powers. One must pick and choose. No matter though; one could throw darts at the pages and be assured of striking something of surpassing interest. In sections I through IV below, I shall nod at each of the main divisions of *Reasons and Persons*. Section V briefly assesses the overall thrust and importance of the

work. Even if this provides a tolerably accurate map of the terrain Parfit covers, it cannot begin to convey a sense of the extraordinary subtlety and brio he brings to these explorations.

I

"For each person, there is one supremely rational ultimate aim: that his life go, for him, as well as possible." (p. 4) So speaks (one version of) the Self-Interest theory of Rationality (S). To it stand opposed other voices, notably those that we identify as demands of morality. While S prescribes ultimate concern for one's own well-being, every nonegoistic version of morality requires at least some sacrifice of self-interest in order to advance the prospects of other persons. Because S and morality are alike in purporting to provide *definitive reasons for action*, the opposition is not benign. Some writers take it to be self-evident that moral reasons must be definitive. The claim of S to be an adequate theory of rationality can therefore be dismissed out of hand. Others take it as evident that S is the definitive theory, and so they either dismiss morality as nonrational or else make it subsidiary to S.[2] A third group, of whom Sidgwick is perhaps the most notable representative, finds the opposition between S and morality insufficient grounds to dismiss either. Sidgwick classifies S as a "method of ethics" because it prescribes over the same range and in the same apodictic tones as does morality. Its reason-giving character is not impugned by its opposition to morality, nor is the reason-giving character of (utilitarian) morality defeated by S. We are left with two mutually inconsistent theories of ultimate and definitive reasons for action. Sidgwick quite appropriately views this as a philosophical failure.

Parfit, like Sidgwick, finds the opposition of S and morality insufficient reason to dismiss either. He does not, however, acquiesce in Sidgwick's reluctant conclusion that the theory of practical reason must remain perpetually schizophrenic. Instead, he adopts the strategy of examining ways in which theories can be *self-defeating*. The lure of S will be resistible if the adoption of S leads to the defeat of the goal that S prescribes as supreme for each agent. This may seem to be a straightforward exercise in demonstrating the internal inconsistency of a theory. Matters are not, however, so simple. Inconsistency is one way in which a theory can be self-defeating, but it is not the only one. Moreover, a theory which, in one respect, is found to be self-defeating may yet survive to prescribe another day.

Parfit's opening jab is to show that S can be *indirectly individually self-defeating*. Kate is a writer whose strongest desire is that her books be as good as possible. Even though, because of the strength of this desire, she sometimes works to the point of exhaustion, writing books provides her much happiness. Were her strongest desire instead that her life go as well for her as possible, she would not drive herself to exhaustion, but neither would she enjoy the same zest for writing. That is, were she to adopt S she would secure less of what S prescribes as the ultimate goal for a person. Similarly, a person with altruistic concerns may derive more satisfaction than someone all of whose desires are self-directed. To generalize, it can be

the case that one who rejects S as determinative for himself will secure more of what S takes to be of ultimate value than he would have accrued had he accepted S.

Does this amount to a refutation of S? Parfit argues that it does not. That is because, roughly, the failure is not that of S but of the agent. It is simply a fact about Kate that she is unable to substitute for her desire to write books some other motivation that will make her life go better. Should Kate's strongest motivation be to act on S then she will, on each occasion, produce the best possible result for herself. However, having that motivation is what makes things go less well for her. Therefore, S counsels that Kate not be motivated always to try to make her life go as well for herself as possible; she does better instead to try above all to write good books. Although Parfit does not put it in this way, we might say that S is properly conceived of as a metaprinciple for the appraisal of first-order principles that guide action. For some persons S will also be an S-optimal first-order principle, but for others it will not.

It is not only S that is indirectly self-defeating. C, Consequentialism, can also prove suboptimal precisely in terms of what C holds to be valuable. For example, persons who aim above all at maximizing the well-being of everyone may be deficient repositories of happiness because they lack ordinary human concerns for specific persons and projects. A world of individuals who succeed perfectly in acting as C directs may contain less happiness than if each had some ultimate motivation other than C.

Parfit next proceeds to consider how theories can be *collectively self-defeating*. When applied to S, this involves a situation in which each person succeeds in doing what is most in his own interest, but where the result for all of them is worse than if each had adopted some other strategy. The vast literature on prisoners' dilemmas provides a stock of examples that Parfit augments with his own ingeniously constructed cases. Prisoners' dilemmas do not, however, demonstrate S to be inadequate. Each participant in a (multi-person) prisoners' dilemma succeeds in bringing about the best possible result for himself. Were anyone altruistic, the outcome for himself would be yet worse. It is each person's bad luck that every other participant is guided by S.

If potential participants in a prisoners' dilemma could choose for themselves and others basic motivations, they would do better to engender morality rather than the belief that S is true. Even if S is the true account of rationality, it recommends, "Believe morality (to be what is most rational) rather than me." S is, says Parfit, *self-effacing*.

So also, he claims, is M, Common-Sense morality. While precise delineation of M is difficult, on any plausible construal it involves special obligations toward persons with whom one stands in relationships of intimate associations. For example, I have a special obligation to care for the children that are mine. I am not at liberty to be entirely cavalier toward other people's children, but I owe them less than I do my own. This can generate suboptimal outcomes. Suppose that I can extend either some minimal benefit to my child or a substantially greater benefit to your child. Suppose also that you are correspondingly situated. If you and I are both guided by

M, we produce a result less M-good for each of us than if we were motivated otherwise. M, in this case is directly self-defeating. Though we succeed perfectly in doing what M requires, we secure less of what M acknowledges to be of value. The flaw is in M rather than in us. Were we in a position to choose basic motivations for ourselves and all others, we do better to instill a motivation to be guided by concern for overall well-being, to move from M closer toward C.

This conclusion seems to me to be questionable. Parfit misconstrues, I believe, the dominant thrust of M. We are, for the most part, *permitted* to lend extra concern to specially regarded persons rather than *required* to do so. And, to the extent that M recognizes a strict obligation in this respect, it is almost always the case that things *will* go less well if individuals gener- ally fail to acknowledge that obligation. Would the goods that we attain through relations to our children and friends be as readily available if we were to change our belief that we owe them special consideration? In some possible worlds, perhaps. But Parfit has not persuaded me that such is the case for the actual world. Also, his characterization of M as self-effacing with respect to C seems reversible. Suppose that you and I come to realize that we will do better by our children if each of us is guided by C. Does this provide us any reason to attempt to alter our volitional makeup? Only by some such consideration as, "By acting to change myself and others I shall make things better for the child who *matters* to me." My decision to be guided by C is based on my prior acceptance of M as a true indicator of where lies value for me.

No bird's eye view of the argument of Part One can convey an adequate sense of the charm of the scenery closer to the ground. Parfit is a better analyst than most, but it is the freewheeling play of his imaginative powers and the daring with which he throws himself into the densest philosophical thickets that excite admiration and delight. In the Anglo-American philosophical world, only Nozick is comparable. Dozens (perhaps hun- dreds—I have not attempted to count) of artfully constructed examples punctuate the text.[3] Usually they illustrate with grace and economy the philosophical issue in question. Occasionally though Parfit is done in by ex- cessive reliance on this talent. For example, he argues that it is rational to take into account even extremely minute probabilities when the stakes are very high. Designers of nuclear reactors should not ignore probabilities as low as one in a million of component failure when the result of such failure will be catastrophe. Similarly, one has good reason to vote in American na- tional elections even if the probability that one's vote will be decisive is on the order of one in a hundred million. The cost to the individual of voting is less than the product of the net average benefit to each American of the superior candidate being elected, the number of Americans, and the prob- ability that one's own vote will be decisive.

If the analogy holds, Parfit has untangled a major problem of normative democratic voting theory. It does not. Nuclear failure is a bad for virtually everyone affected and bad for each in nearly the same way. Political goods and bads are notoriously less clear-cut, and predictions about what can- didates will do if elected are extremely tenuous. If a voter guesses incor- rectly, he "cancels out" the vote of someone who has guessed better. There- fore, consequentialist rationality obliges one to apply to one's cost-benefit calculations an "epistemic rate of discount." In particular, one does better *not to vote* if one's Political Judgment Quotient ranks below the median of all

prospective voters. If each prospective voter applied this consideration and could be confident that all others would do so, the outcome that emerges as optimal is not *everyone* voting but only *one person* voting—the Philosopher King redivivus! This may seem too quick to the reader, and indeed it is. But it identifies more of the relevant parameters than does *Reasons and Persons*. The moral is that even a uniquely gifted philosophical fabulist must not overly burden his muse.

A notable corollary of the investigation of collectively self-defeating action is Parfit's demonstration of the ungainliness of ordinary moral reasoning when applied to situations in which one's actions affect each of many other people to a negligible degree. A million polluters collectively impose severe harms on the populace, but no one polluter imposes a perceptible harm on anyone. Parfit argues persuasively that our moral principles have largely emerged in response to the experience of readily identifiable harms and benefits being visited on readily identifiable individuals. They do not transfer easily to large number cases. Yet the importance of large number cases under conditions of contemporary urban life is profound. Unless we radically revise our moral architecture, the results are liable to be very bad. Parfit contends that our moral reasoning must give weight to imperceptible effects of action and to what people *together bring about*. While I have qualms concerning some of the specifics of the diagnosis, the discussion undeniably enriches moral inquiry.

II

In Part Two, Parfit resumes the attack on S. S mandates temporal neutrality; a person acts rationally only if he gives equal weight to every state of his life. Against this Parfit presents what he calls the Present-aim theory, P. According to P, one does best to act to achieve one's present desires. P can take various forms, and the one Parfit commends is the Critical Present-aim theory, CP. According to CP, one does best to act to achieve those of one's present desires that are intrinsically rational. There is an uninteresting way in which CP can collapse into S: one's dominant present aim may be that things go as well as possible for oneself over the course of a lifetime. Another possibility is that it is rationally obligatory to value every period of one's life. This version of CP is not, however, equivalent to S. One may rationally value some periods above others, and one may value things other than one's own well-being. For example, I may be willing to sacrifice on behalf of my friends or my projects. If so, I cannot thereby be condemned as irrational.

The S theorist has a difficult time contending against CP. He can point out that one who always acts to advance his present desires can bring it about that future desires will be less well satisfied. On each occasion one succeeds in advancing present desires, but they are less well satisfied in toto than if one had adopted S as one's policy. That is to say, CP is indirectly self-defeating. But, as we have previously seen, so too is S. If that does not constitute a decisive objection to S, then neither does it disable CP. The S theorist can maintain that to lend weight to only present desires is rational only if those are the only desires that will be one's own. Since one will come to have other desires, one cannot rationally decline to give them weight in the present. But if the complaint is that partiality toward the present is irrational, how can partiality toward oneself be justifiable? Not all desires are *present* desires, but neither are all desires *my* desires. S stands uneasily be-

tween CP and C. By artfully playing the two strings to his bow, Parfit unmercifully harries S. What may have initially seemed to be the obviously correct account of practical rationality emerges battered and bruised.

Has Parfit successfully solved Sidgwick's problem? I believe that he has. However, I do not believe that the problem was ever as deep as Sidgwick took it to be. When the S theorist maintains, "For each person, there is one supremely rational ultimate aim: that his life go, for him, as well as possible," what is the force of the qualifier "for him?" It may mean that I am not to count the outcome of a project of mine as successful unless the beneficiary is me. The only ends at which one rationally can aim are self-referential. So understood, S is palpably false. *Of course* I am rationally entitled to aim at goods that are not uniquely goods for me. (That is why a desire-fulfillment account of practical rationality is more plausible on the face of it than a narrowly hedonistic theory.) Alternatively, the qualifier may be otiose. On this construal, one does well to lead a life in which one acts to achieve what one takes to be valuable, whether or not that value resides within the confines of the self. For example, I can succeed in living a life devoted to the production of beauty even though it is not myself that I render beautiful. So interpreted, S becomes difficult to reject. But then there is no occasion to reject S; it is compatible with any credible theory of what constitutes a well-lived life. The crucial ethical questions become: (1) What lives are well-lived for beings such as ourselves?; and (2) What principles ought we acknowledge as applicable if we are to live such lives? These remain difficult issues of the utmost importance, but what need no longer detain us is the arid question of how to square S with morality.

Parfit apparently agrees. He says:

> [C]onsider artists, composers, architects, writers, or creators of any other kind. These people may strongly want their creation to be as good as possible. Their strongest desire may be to produce a master-piece, in paint, music, stone, or words. And scientists or philosophers, may strongly want to make some fundamental discovery, or intellectual advance. These desires are no less rational than the bias in one's own favour. (p. 133)

However, his major assault on S is waged not in terms such as these but through the consideration of conundrums involving temporal location. S dictates temporal neutrality. Does that mean that I should care as much about events that *have happened* as I do about events that *will happen*? Suppose that I have been given a drug that affects memory, and I cannot remember whether a painful operation I must undergo has already occurred or whether I will experience the agony today. Should I be neutral between these two possibilities? Or suppose that I have now totally and irrevocably lost some desire that consumed me over much of my earlier life. Am I rationally required to lend weight to its fulfillment proportionate to its duration and intensity? Am I required to assign it *any* weight in current deliberations over what I shall do? The obvious answer is that there is a complete asymmetry between the past, on the one hand, and the present and future on the other. It will be difficult for the S theorist to account for the asymmetry in other than a question-begging manner. Moreover, the asymmetry seems to affect the first-person perspective in a way strikingly different from the third-

person perspective. Suppose I have been informed that my terminally ill mother will suffer great pain for several days before she dies. I am disconcerted by the prospect of her future misery. Then I receive a corrected report; her suffering and death have already occurred. Coming to find out that *her* pain is in the past does not bring the relief from distress that the corresponding discovery about *my* paid did. Through these and a host of related constructions, Parfit raises (and illuminates) many puzzles about time, the concept of prudence, and related issues. I do not mean to suggest anything other than the greatest admiration for this section of the book when I say that it is only of secondary importance in its avowed aim of relaxing this grip on us of S.

III

No one has contributed more to the contemporary discussion of personal identity than Parfit. Previous investigators had divided on whether criteria of identity are best to be thought of in terms of physical or psychological continuities. Parfit has shown that both approaches are vulnerable in essentially equivalent ways. Continuity, whether physical or psychological, is a matter of degree, while identity is all-or-nothing. That is, I may share many or few memories and character traits with some future person; the brain of that future person may have been constructed as a replica of mine or may possess any percentage from 0 to 100 of the cells that are now in my brain. But that future person either is me or is someone other than me. If it is me, then I have survived. If it is merely someone closely resembling me, then I am defunct. Between my survival and my demise there is all the difference in the world. Or so we commonly believe. Parfit's major point is that this is a mistake. Identity is all-or-nothing, but what properly matters to us are all matters of degree. Philosophers can, of course, work at developing acceptable criteria for identity over time, as does Parfit himself. But they should not suppose that they are thereby investigating that which is most important about ourselves.

Those familiar with Parfit's previous work on personal identity will find much to admire in these pages but little that surprises. Fission and fusion of persons, split-brain phenomena, teletransportation, and transplants: all the old cast are back and put through their paces again. Interesting new responses to discussions by Bernard Williams and Thomas Nagel are included, but most of this section replicates earlier essays. It is perhaps fitting that an author who argues that psychological connectedness and continuity are what properly matter to persons has constructed an essay so connected to what he has offered before! This is not intended as criticism; Parfit has established a research program of major importance that he here advances with inexorable thoroughness and energy.

It is likely that considerations of *Reasons and Persons* will concentrate on this section. If so, that will be unfortunate. The metaphysics of personhood merits further attention but so too do the numerous other lines of inquiry opened up by this book. Above all, this is a pathbreaking work in *moral philosophy,* and recognition of it as such should not be deflected by the undeniable charms of its metaphysical excursions. Parfit explicitly places

his discussion of personal identity in the context of what we should care about and how we should act to give effect to that care. He deserves to be taken at his word.

Parfit describes his position as *Reductionist*. Identity is constituted by relations of psychological and physical connectedness and continuity. (State A is *connected* to state B to the extent that they share psychological or physical components. If A is connected to B and B to C, then A is *continuous* with C. That is, connectedness is not a transitive relationship, while continuity is.) Of the two, psychological relationships are of much greater importance than physical ones.[4] There is no "deep further fact," such as a Cartesian ego, in which identity resides. A corollary of this position is that those questions of great practical import that we characteristically phrase in terms of identity are more usefully considered in terms of the constitutive factors of continuity and, especially, connectedness. What I properly have reason to be concerned about is whether, at some future time, there will exist some person connected or continuous with me, not whether that individual will be numerically identical with me. Whether moral obligations engendered by promises and contracts remain in force depends not on whether the promiser survives but on whether there exists a person psychologically connected to a sufficiently close degree with the promiser. Relations to future stages of oneself should be viewed more on the model of relations to other persons. So, for example, paternalistic intervention is moved from the purely self-regarding realm, into which moral considerations are forbidden entry, and placed within the domain in which principles function that mandate permissible treatment toward other persons. The abortion debate becomes recast once we take the important relationship between fetuses and subsequent persons as a matter of degree rather than identity. And so on.

I shall refrain from discussing the many problem cases Parfit constructs on the way to his reductionist conclusion. Surely they will receive extensive examination elsewhere. Even if some details turn out to be in need of reformulation, it is unlikely that reductionism itself will run into much opposition. Cartesian egos are not, after all, much in fashion these days. And if they were, Parfit's steamroller would have sufficed to lay the ghost to rest. Instead, I wish to briefly question whether Parfit has gotten the normative implications right. If he has, one simple conclusion follows: moral philosophy must be rebuilt from the ground up. At least within the Western liberal tradition, the divide between distinct persons is taken to be fundamental. Propositions about rights and justice rest on it. If what separates numerically distinct persons is instead relatively shallow or, more precisely, if it is approximately as deep/shallow as what separates temporally distant stages of one individual's life, then most of what we say about justice, about rights, about respect for persons is insupportable. Secular ethics minus robustly distinct individuals is as impoverished as theological ethics without God.

There are two ways in which one can take personal identity over time to be a deep fact, a fact that undergirds ethics. First, one can believe identity to be a metaphysical given. The world just breaks down into separate persons who remain the selfsame beings from birth (or conception) to death. This metaphysical fact has numerous significant practical implications, and it is the job of normative ethics to spell out what they are. Parfit has, I shall assume, shown this approach to be unsustainable.

Second, one can, as it were, begin where Parfit leaves off. Identity as an

externally conferred fact is shallow. It is not congruent with that which we have reason to value. The concept of personal identity that matters to the metaphysican turns out not to matter for the ethicist. However, the concept of personal identity that figures prominently in developmental psychology and philosophical anthropology is of the first importance. It is the identity which individuals forge over time through their attachment to special concerns and projects that subsequently have directive force over their lives, that determines what can and cannot figure as potential items of value for them. That robust identity in this sense will obtain is not guaranteed by an external blank check signed by Descartes's God. Individuals, though, have an interest in constructing through their own activities lives that are coherent. Identity so construed is in the *active* rather than *passive* mode. And it matters a great deal. Or rather, it matters a great deal if anything does. For suppose that what holds out abiding value to me is end E. Then I have a fundamental interest in acting to realize E and, thus, in constructing a life coherently regulated by the pursuit of E. My success presupposes continuance as an active being whose identity is a function of his directive concerns. Continuance as a Cartesian cipher is irrelevant.

Much follows from this. Individuals who have reason to value the construction of a life coherent in virtue of the persistence of projects have reason to value the liberty to do so. Requirements of noninterference will be grounded in a fundamental separateness of persons that is not passively conferred but actively developed. Acceptable moral principles will recognize that each person has reason to be partial to the ends that are distinctively his own. Rights that secure to each person a privileged moral space within which his own will is sovereign manifest this recognition. Paternalism is suspect not because future selves are assured of being closely connected to present ones but for nearly the opposite reason. I may be the chief creator or the shape that my future self takes or this job may be taken out of my hands by would-be benefactors. Even if the benefaction is genuine, my interest in constructing *from the inside* the life that is mine has been abridged.

The two preceding paragraphs are perhaps too schematic and abridged to be persuasive. Certainly, a much expanded treatment is called for if we are adequately to assay what it is about the identity of persons that is important for moral philosophy. Still, I hope to have rendered it at least credible that traditional ethics very nicely survives Parfit's reductionism about persons. Actually, I believe a stronger result to obtain. Parfit sometimes presents his account of persons as above all threatening to classical liberalism. Collectivist principles, he suggests, are to be substituted for individualistic ones. Here, I think, Parfit badly misappraises the significance of his own work. Classical liberal views not only can survive on a diet of Parfitian reductionism but, in fact, are better nourished than ever they could be on a fare of Cartesian egos whose identity over time is a "deep fact." A sort of division of labor is to be observed; because metaphysics cannot read much *out* of the identity of persons, practical philosophy reads much *into* it.

I cannot quite leave the matter there because Parfit does not. His "official" view is that reductionism is revisionary of fundamental moral attitudes. (It also strikes yet another blow against S.) But in the closing pages of the book he notes:

> On the Non-Reductionist View, the deep unity of each life is automatically ensured, however randomly, short-sightedly, and

passively this life is lived. On the Reductionist View, the unity of our lives is a matter of degree, and is something that we can affect. We may want our lives to have a greater unity, in the way that an artist may want to create a unified work. And we can *give* our lives greater unity, in ways that express or fulfil our particular values and beliefs.

It may appear then that there is no substantive disagreement between Parfit and myself, simply a difference in how widely each of us will employ the term "identity." That is not so; there remains a wide disparity between our appraisals of the implications of reductionism for moral theory. Parfit's revisionistic conclusions are resistable, and in the cited passage, he takes the crucial first step toward doing so. It is indicative of his extraordinarily wide vision and philosophical *fairness* that he offers an observation potentially undermining of theses he has labored to bring forth and defend. I suspect that, in his subsequent work, he will move yet further from the official position.

IV

How we act affects the welfare of persons. How we act also affects which persons there will be. The former proposition has been at the forefront of ethical theory forever; the latter has hardly been noticed. Section IV goes a long way toward righting the balance. If conceptions were to occur between different partners, or more than a few days later or earlier than they in fact do, then the resulting persons would be different persons. Major social or economic policies influence who meets whom, whether a couple conceives a child, and when they do so. Thus they affect which people there will be. The point, once stated, is obvious enough. To the best of my knowledge though, prior to 1970 its implications were entirely unexplored.[5] And that is a surprising omission because, as Parfit unmistakably demonstrates, the implications are anything but trivial.

Suppose that a couple could conceive a child this month who, predictably, would be severely handicapped, though with a life worth leading. They could instead wait another month to conceive, in which case their child would enjoy normal prospects. Should they do the former, it seems to be a choice that is patently wrong. Yet to whom has the wrong been done? Not to themselves, for even if they have acted imprudently, that's not the mistake we wish to spotlight. Neither is it a wrong done to the child who comes to be, for *that* child is benefited by their decision to conceive earlier. It would not exist, and thus enjoy the life we acknowledge to be worth living, had they waited. Finally, it is metaphysically confused to locate the wrong in the child who would have existed had they conceived one month later. We may want the scope of our moral principles to be generous, but to extend them to nonexistent beings is profligate. Either we withdraw the claim that their act was wrong, or else we recognize it as an act that is not a wrong *to* anyone. But what principle or principles do we invoke to justify that kind of judgment?

Similar problems arise at the macro level. Suppose that the social choice is between two different schedules of depletion of natural resources. One is *steady-state,* in which each generation will predictably enjoy a stock of

resources equal to that of its predecessor, technological advance generating new resource availability at roughly the same rate at which old ones are expended. The other policy is *rapid depletion*. For three centuries resources will be expended at a rate that maximizes current economic prosperity. People live happier lives than they do under steady-state. However, at the end of three centuries quality of life will plummet, though not below a point at which it becomes not worth living. Which policy ought to be adopted? Before a response is hazarded, it should be noted that any decision about economic policy will significantly affect patterns of conceptions. It is almost certain that, if rapid depletion is adopted, there will not exist, three hundred years later, any individual who would have existed had steady-state been adopted. People will be worse off, but they will be *different* people. No one will be able to complain that the choice of policy has rendered him worse off. On the other hand, if the choice is steady-state a large number of people will be worse off than they might have been. That will be the case for all members of the current generation and diminishing numbers from subsequent generations. If we apply only person-regarding principles of choice, rapid depletion wins. Yet that seems to be the wrong result.

Such considerations are the takeoff point for Parfit's investigations. The goal is the discovery of a set of impersonal moral principles, Theory X, that yields some reasonable conclusions and no absurd ones. Modest enough, one might think, yet the quest turns out to be exceedingly difficult. Suppose that one adopts the classical utilitarian principle of maximizing total happiness. Then we should reject a world containing ten billion very happy people in favor of bringing about a world containing a vastly greater number of people each of whom leads a life barely above the threshhold at which life is worth living. Maximizing average happiness does no better, for then we would be required to bring about a world containing only a handful of ecstatics rather than one in which billions of people lead very fulfilling lives. An intermediary position is that quality and quantity both matter, but that quantity matters less and less as the numbers grow. (Actually Parfit identifies several distinct variants of the intermediate position.) This receives a good run for its money but is found to lead to other unacceptable results. One interesting reason why it does so is because there exists an asymmetry between happiness and unhappiness. The intermediate position takes it to be a matter of near-indifference whether the world brought into being contains a very large number of very happy people or a world just like it except twice as populous. Our intuitions do not rise in opposition. Yet it is not a matter of indifference whether we act to bring about a world of many beings who lead lives of excruciating agony or instead a world just like it except with twice as many people suffering. The value of increments to the quantity of happiness asymptotically approaches zero, but the disvalue of incremental suffering does not.

It is not possible to chart here the many byways that follow. In this section, as elsewhere, Parfit aims at nothing less than examining *all* feasible alternative views. If he does not succeed, that is because *no one* could, not because of any discernible philosophical omission on his part. Still, the quest fails. Each reshaping of principles yields paradoxical results of its own or else can be shown to entail results that have previously been dismissed. Theory X remains elusive. It may be altogether unobtainable. That will be the case if we demand consistency with our pretheoretic intuitions, but those

intuitions cannot themselves be consistently expressed. There may be some comfort for the theorist in this; the fault will not be in him but in the stars (that deceptively twinkle in people's minds). Unfortunately, this is not a dilemma purely of theoretical interest. Parfit is absolutely persuasive in arguing that we possess unprecedented power profoundly to affect the well-being of future generations. This is a genuine novelty in human history. We do not enjoy the luxury of being at liberty to confine our moral speculations to the range of cases familiar since Aristotle. New potentialities call for theoretical rearmament. What lies ahead of us is vastly more important than the totality of what lies behind: "The Earth will remain habitable for at least another billion years. Civilization began only a few thousand years ago. If we do not destroy mankind, these few thousand years may be only a tiny fraction of the whole of civilized human history." (pp. 453-454) Parfit has not found Theory X but, almost certainly, he has opened up what is bound to become a major area of moral inquiry.

V

What constitutes the unifying thread of this massive work? Parfit tells us, "My two subjects are reasons and persons. I have argued that, in various ways, our reasons for acting should become *more impersonal* (p. 443, emphasis Parfit's). This accurately characterizes much of the text, most especially Part Four. The move from S to C is from personal to impersonal reasons but, as Parfit himself notes (p. 445), CP is, in some ways, *more* personal than S. I have already discussed the extent to which Parfit's reductionism about personal identity entails impersonalism. If I am correct, reductionism is hospitable to recognition of personal reasons. Parfit, in one of his voices, seems to agree. So Parfit's own characterization of the upshot of his text is, at least in part, misleading. Does this matter?

Not very much I think. Tight thematic unity is the primary virtue of some books. Even a narrow range of ideas has some considerable effect if they are kept carefully in line. *Reasons and Persons* though is bursting with ideas. Philosophical discussion of each of the major areas addressed by Parfit will be transformed in the wake of this book. Even some peripheral forays make important additions to the relevant literature, e.g., the treatment of prisoners' dilemmas. Rarely in contemporary philosophy has such breadth of ambition been harnessed to such depth of achievement. That is probably too cautious; without much danger I could remove the qualifier, "contemporary." I noted in section I Parfit's inheritance of concerns from Sidgwick. He explicitly voices his great admiration of Sidgwick and, though he nowhere says so, it is clear that he wants to write a book that can stand shoulder to shoulder with *Methods of Ethics*. I think he has done so. *Reasons and Persons* lacks the full systematic coherence of the earlier book but compensates in virtue of its imaginative richness and the unconstrained delight in doing philosophy that springs from the pages.

I do not wish to be perceived as a *total* enthusiast. (But why not?) This is not a book without flaws. Parfit is quick to conjure illuminating examples but, as noted above, occasionally an excess of confidence in his way with an analogy leads him astray. And beyond citation of particular instances, a question that goes insufficiently examined here is that of the proper place of

fanciful thought experiments within philosophy. Parfit offers some useful justificatory remarks but does not consider whether science fiction scenarios might be of far greater utility in metaphysics than in moral philosophy. The metaphysician's eye ranges over all possible worlds; the moral philosopher has a difficult enough time prescribing for the actual world and its close neighbors. Given concerns rooted in *practice*, it need not be an objection to a moral theory that it stands mute in the face of people fissioning and fusing, that it fails to tell us how we ought to act if each person's intentions were transparent to all others, and the like.

Despite its analytical precision, the book gives the appearance of having been rushed. Some thoughts appear to be second thoughts, not fitting comfortably with the official position Parfit espouses. Further pruning might have eliminated some tensions in the text. This though may not be fair to Parfit. The difficulty may be less too quick a rush to publication than a laudable unwillingness on the part of the author to stem the onrush of ideas. Some philosophers make a deliberate policy of sequestering their arguments from all considerations that might render them uncomfortable. Parfit is a philosopher both honest and forthcoming. It is hard to imagine him engaging in philosophical protectionism. Less controversially, one wishes that this otherwise handsomely printed volume had received more attentive proofreading. The sentence running between pages 39 and 40 is hopelessly garbled; one or more lines are missing on page 193; a middle paragraph on page 331 is unintelligible; the footnote numbered "2" on page 448 should be "1"; footnote 6 on page 523 is incoherent. A number of other typographical missteps less damaging to the readability of the text are also present. These are quibbles, but when philosophy is done so well, one does not want even one sentence mutilated in the passage from author fo publisher.

LOREN LOMASKY

University of Minnesota, Duluth

1. See "Personal Identity," *Philosophical Review* 82 (1971): 3-27; "On 'The Importance of Self-Identity'," *Journal of Philosophy* 68 (1971): 683-690.
2. Neoclassical economic theory characteristically makes it a matter of definition that S provides the correct account of rationality.
3. In an autobiographical aside (p. 157), Parfit mentions that between the ages of 7 and 24 what he most wanted was to be a poet. Despite his declaration to the contrary, I am not convinced that he subsequently changed his mind.
4. Parfit is prepared to admit (p. 284) that this may not be so for very beautiful people!
5. Parfit raises the issue in "On Doing the Best for Our Children," in *Ethics and Population*, ed. Michael Bayles (Cambridge: Schenkman, 1976), pp. 100-115. It received independent statement in a different context in Robert M. Adams, "Existence, Self-interest, and the Problem of Evil," *Nous* 13 (1979): 53-65.