Oxford Handbooks Online

Abstract and Keywords

Perhaps the most familiar form of consequentialism is classical hedonistic act utilitarianism, which claims, roughly, that an agent ought to perform that action, among the available alternatives, that produces the most net pleasure (pleasure, less pain) for everyone concerned. But this classical form of utilitarianism is thought by many to be just a special case of a more general or abstract class of consequentialist moral theories that make the moral assessment of alternatives depend in some way upon their value. How to understand and assess consequentialism depends on how one specifies this more general class of theories. This article understands consequentialism quite broadly, with the result that it is a large and heterogeneous family. This makes it difficult to get very far discussing the prospects for consequentialism as such. Different varieties of consequentialism have different strengths and weaknesses.

Keywords: consequentialism, classical hedonistic act, utilitarianism, moral theories, moral assessment

Perhaps the most familiar form of consequentialism is classical hedonistic act utilitarianism, which claims, roughly, that an agent ought to perform that action, among the available alternatives, that produces the most net pleasure (pleasure, less pain) for everyone concerned. But this classical form of utilitarianism is thought by many to be just a special case of a more general or abstract class of consequentialist moral theories that make the moral assessment of alternatives depend in some way upon their value. How to understand and assess consequentialism depends on how one specifies this more general class of theories. I will understand consequentialism quite broadly, with the result that it is a large and heterogeneous family. This makes it difficult to get very far discussing the prospects for consequentialism as such. Different varieties of consequentialism will have different strengths and weaknesses. Of necessity, my discussion will be selective, concentrating on those varieties that seem to me to have a significant tradition or to be especially interesting.

1. Consequentialist Structure and Varieties

Classical hedonistic utilitarianism conceives of the good in terms of pleasure and identifies an agent's duty with his promoting pleasure. This makes the good explanatorily prior to the right, insofar as it defines right action in terms of promoting the good (see Rawls, 1971, sec. 5). Generalizing, we might understand consequentialism as the set of moral theories that make the good explanatorily primary, explaining other moral notions, such as duty or virtue, in terms of promoting value. For instance, a consequentialist conception of duty might identify an agent's duty as an action that promotes the good, whereas a consequentialist conception of virtue might identify virtuous dispositions as those with good consequences. We can construct consequentialist analyses of virtually any object of moral assessment, including actions, motives, individual lives, institutions, and moral codes. To be a consequentialist about the assessment of any of these things is to think that one's assessment of alternatives within that domain should be governed in a suitable way by the comparative value of the alternatives. Understanding consequentialism this way equates it with a teleological conception of ethics.

But consequentialism, as such, is neutral about a great many issues. To make discussion manageable, it will help to focus on one kind of consequentialist analysis. One traditional focus concerns the analysis of *duty* or *right action*. Many issues that arise in understanding and assessing consequentialist conceptions of right action apply *mutatis mutandis* to other kinds of consequentialist analysis.

Consequentialism takes the good to be primary and identifies right action as action that promotes value. As such, it contrasts with two different conceptions of right action. *Deontology* takes right action to be the primary evaluative notion; it recognizes various actions as obligatory, prohibited, or permitted on the basis of their intrinsic natures and independently of the value they produce. *Virtue ethics* takes the idea of a morally good character to be explanatorily primary in the account of right action; right action, on this view, is action performed by someone with a virtuous character or that expresses a virtuous character.[1]

The consequentialist conception of right action leaves several questions unanswered. One pertinent question concerns *what* is valuable. This is a question about what has intrinsic value. It is, in part, a question about the human good. What are the constituents of a good human life? One familiar conception is the *hedonistic* claim that pleasure is the one and only intrinsic good and that pain is the one and only intrinsic evil. Alternatively, one might understand the human good in *preference-satisfaction* terms, as consisting in the satisfaction of actual or suitably informed or idealized desire. Hedonism and preference-satisfaction views (p. 382) construe the human good as consisting in or depending upon an individual's contingent and variable psychological states. By contrast, one might understand the good in more objective terms, either as consisting in the *perfection* of one's essential capacities (e.g., one's rational or deliberative capacities) or as consisting

in some *list* of disparate objective goods (e.g., knowledge, beauty, achievement, friendship, or equality).

Connected with these issues are other questions about the good. Who are the *bearers* of intrinsic good? Are there goods for all sentient creatures, or only for a more limited class of beings, such as human beings or persons?

Another question is whether some things are valuable independently of the contribution that they make to the lives of sentient creatures. If so, we might say that there are *impersonal* goods. Some people think that beauty and equality are impersonal goods. But even if they are intrinsic goods, it is debatable whether they have value independently of any contribution they make to sentient life. If there are no impersonal goods, we might say that all goods are *personal* or *sentient*.

Still another question is *Whose value matters?* Should an agent be concerned about all those that it is within her power to benefit, and among those that demand her concern, should they matter equally? At one extreme lies the *impartial* consequentialist view that an agent should be concerned to promote any and all kinds of value and, in particular, should have an equal concern to promote the well-being of all those that it is in her power to affect for better or worse. Utilitarianism is probably the most familiar form of impartial consequentialism. It instructs agents to promote human or sentient happiness generally. But a view that recognized impersonal values and instructed agents to promote these wherever possible would also be a form of impartial consequentialism. At the other extreme lies the *partial* consequentialist view that an agent should be intrinsically concerned with promoting only her own welfare. Such a view would be a form of ethical egoism. In between these extremes lie more *moderate* forms of consequentialism that demand intrinsic concern for others but that limit the scope or weight of such concern. One example of such a moderate view is the view that C. D. Broad called "self-referential altruism" and associated with common-sense morality (1971, esp. p. 279). Self-referential altruism claims that an agent's concerns should have wide scope, but variable weight. It says that an agent has an obligation to be concerned about anyone that it is in her power to benefit but that the weight of an agent's moral reasons is a function of the nature of the relationship in which the agent stands to potential beneficiaries. On this view, an agent has reason to be concerned for perfect strangers as well as intimate associates, but, all else being equal, she has more reason to be concerned about the well-being of an associate than a stranger.

These distinctions within consequentialism can also be made in terms of the distinction between agent-neutral and agent-relative reasons. The general form of ^(p. 383) *agent-relative* reasons makes essential reference to the agent in some way, whereas the general form of *agent-neutral* reasons does not (see Nagel, 1986, p. 152). Being under a duty to help children, as such, would involve an agent-neutral reason, whereas being under a duty to help one's own children would involve an agent-relative reason. Being under a duty to minimize suffering would be an agent-neutral reason, whereas the deontological duty never to be the cause of another's suffering, even if this is necessary to minimize

total suffering, would be an agent-relative reason. Some writers have believed that this distinction between agent-neutral and agent-relative reasons is the way to distinguish between consequentialist and nonconsequentialist moral conceptions. But I do not share this view. Whereas only utilitarianism and other forms of impartial consequentialism will qualify as agent-neutral conceptions, some agent-relative conceptions are consequentialist. Ethical egoism and self-referential altruism both identify an agent's duty with promoting values, though they limit the scope or vary the weight of the values she ought to promote.

Consequentialists are concerned to promote the relevant values. This contrasts with the deontological response to value. To *honor* a value is to act on it or protect it at every opportunity. To *promote* a value is to take steps that lead to its greater realization overall. But promoting a value overall can require failing to honor it on some occasions, as it would, for example, if promoting and protecting freedom within a community required establishing a compulsory draft. And honoring a value on some occasion may involve failing to promote that value, as it would, for example, if saving an innocent life now could only be done in ways that prevented saving even more innocent lives at some later point in time. Whereas the consequentialist tells agents to promote the relevant values, the deontologist tells them to honor those values (see Pettit, 1991).

There are different ways of promoting values. Some ways of promoting values are *direct*, inasmuch as they assess alternative actions by the contribution that each alternative makes to the relevant values. The most traditional direct form of consequentialism is *act* consequentialism, which says that an agent should perform that action whose value (of the relevant sort) is at least as great as that of any alternative available to her (or at least one such action, if there are multiple actions meeting this condition). Act consequentialism tells the agent that it is her duty to maximize value. Some have found this act consequentialist claim too burdensome. In requiring agents always to do the best, act consequentialism seems unable to accommodate the idea of supererogatory actions—those actions that in some sense go beyond or are better than what is required by duty. Impressed by this worry, some direct consequentialists have looked for less demanding ways of promoting value. One such view is a *scalar* consequentialism. On this view, one alternative is morally better than another if it produces more of the relevant kind of value and morally worse if it produces less. The scalar view, as such, does not (p. 384) say what an agent's duty is. The scalar view is sometimes advanced as part of a *satisficing* view. The satisficer demands of the agent, not that she maximize value (the relevant values), but rather that she perform any of the alternatives that are good enough—that is, that lie above some specified threshold of value. Duty only requires that the agent perform an action above the relevant threshold. If she chooses an action far above the threshold, for instance, one that is at the top of the scale and maximizes the relevant values, then she has gone beyond her duty and done something supererogatory (see Slote, 1985, chs. 3, 5).

As we will see, there are other sorts of concerns about direct consequentialism. These are traditionally formulated about act or maximizing consequentialism, but I think that they apply, with suitable changes, to satisficing forms of consequentialism too. Roughly, the worry is that maximizing values will sometimes require agents to deviate from moral precepts that seem independently compelling. Perhaps honesty is generally the best policy for both egoists and utilitarians, but there must be cases in which the agent is better off or humanity is better off if the agent is dishonest. If honesty is an absolute moral demand, this spells trouble for direct consequentialism. And even if honesty is not an absolute demand, it may seem to be a more robust demand—less sensitive to consequences—than direct consequentialism would imply. Concerns such as these have led some to endorse *indirect* forms of consequentialism that assess actions not in terms of their values but rather in terms of the value of the rules or motives under which the action can be subsumed. So, for example, rule utilitarianism claims that an action is right just in case it conforms to a rule the general acceptance of which by humanity would have consequences at least as good for humanity as any alternative rule (see Brandt, 1963; Hooker, 2000). Rule egoism would say that an action is right just in case it conforms to a rule the general acceptance of which by the agent would promote his welfare at least as well as any alternative rule available to him. Just as rule consequentialisms identify duty with acting on optimal rules, motive consequentialisms identify duty with acting on optimal motives (see Adams, 1976; Gauthier, 1986). These forms of indirect consequentialism will be responsive to worries about direct consequentialism, insofar as the motives and rules recognized by common-sense morality have optimal acceptance value.

We can see how consequentialism, so conceived, forms a large and heterogeneous family of moral theories. Though some generalizations about consequentialism are more robust than others, it is difficult to assess the strengths and weaknesses of such a disparate set of claims. It will make discussion more profitable to focus on a few main forms of consequentialism.

(p. 385) 2. Maximization

First, I shall focus on maximizing versions of consequentialism. It will be easier to understand this focus within the framework of direct consequentialism, even if the debate between maximizers and satisficers cuts across the debate between direct and indirect consequentialism. One problem for the pure satisficer is that she seems to have no basis for choosing among or ranking options all of which are above the threshold of permissibility. If we discriminate only between options above and below the threshold, then it seems a matter of indifference how far above the threshold one is. But that is counterintuitive. Why should the value of options matter just up to the threshold and not at all above it? Indeed, the pure satisficer has a problem explaining why the best is typically supererogatory and deserving special praise. But these objections dissolve if satisficing is combined with the scalar view. For the scalar part of the view allows one to make moral discriminations among all options, both below and above the threshold, that

Subscriber: King's College London; date: 14 February 2019

track their value; the satisficing part of the view says that options below the threshold are impermissible and that all options above the threshold are permissible. So the scalar-satisficing view allows us to say why the best is supererogatory and deserves praise.

The main rationale for satisficing is that maximization seems too demanding. Performing the best option is typically permissible and admirable. But we might be reluctant to say that it is one's duty. For that would imply that all suboptimal actions, even very good ones, are wrong. And many such actions we would be reluctant to blame the agent for performing. It is tempting to say that duty only requires performing some action above a certain value threshold, that any action above the threshold is permissible, and that the very best action is, at least typically, supererogatory rather than obligatory. The scalar-satisficing view respects these intuitions. By contrast, maximizing act consequentialism seems to violate them, inasmuch it seems to imply that the optimal action is always obligatory, that all other actions, however good, are impermissible, and that there is no such thing as supererogatory action.

One reason the maximization may seem harsh is that it seems to require that we blame the agent for every suboptimal act, however good. But this does not follow. The maximizer must assess actions and responses to those actions separately. Even if suboptimal acts are wrong, it doesn't follow that it's good to blame them. They may be cases of *blameless wrongdoing* (see Parfit, 1984, ch. 1). Indeed, if sufficiently good, suboptimal actions not only need not be blameworthy but are likely to be praiseworthy. These would be cases of *praiseworthy wrongdoing*. These observations suggest a way in which the maximizer might try to capture the intuitions to which the scalar-satisficer appeals. Common-sense morality distinguishes among (1) the obligatory, (2) the permissible, and (3) the supereroga (p. 386) tory. Though the maximizer makes the optimal obligatory, treats all suboptimal acts as impermissible, and does not strictly recognize actions that are morally better than one's duty, he can nonetheless draw a similar tripartite distinction among (1a) acts whose omission is blameworthy, (2a) acts whose omission is not blameworthy, and (2c) acts whose omission is not blameworthy and whose performance is praiseworthy (or perhaps deserving of special recognition and praise). Of course, these notions of praiseworthiness and blameworthiness must themselves be interpreted in maximizing consequentialist terms. While there is no a priori guarantee that the maximizer's tripartition will track perfectly the common-sense tripartition, there is reason to think that they will sort options in similar ways and to wonder whether the maximizer's tripartition might not provide reflectively acceptable guidance and correction where the common-sense tripartition provides uncertain or questionable guidance.[2]

Maximizing consequentialism is the more traditional form of that doctrine. Because it is not clear that its scalar-satisficing rival enjoys any real advantages, it will be simpler to focus on the traditional conception. However, much of what I say about the traditional conception applies, *mutatits mutandis*, to satisficing conceptions.

---

3. Direct Consequentialism

Second, I shall focus on direct, rather indirect, forms of consequentialism. Direct consequentialism assesses all things, including actions, in terms of the value of their consequences, whereas indirect consequentialisms assess actions in terms of their conformity to rules, motives, or dispositions with good or optimal acceptance value.

Indirect forms of consequentialism are worth discussing separately only if they have different implications from direct consequentialism. Will actions conforming to motives or rules with optimal acceptance value be different than the optimal actions? If we notice that motives can be very discriminating and rules can be fine-grained, we might wonder if the best motives and rules wouldn't always require the best actions. But it is often difficult to identify reliably the optimal action, and an agent may often do better overall by internalizing and acting on some fairly coarse-grained set of motives and rules than by attempting to optimize in each of her actions, even if this means performing some actions that are suboptimal. If so, it seems that acting on the best rules or with the best motives might not be the same as performing the best actions.

Even if indirect consequentialism is a genuine alternative to direct conse (p. 387) quentialism, we may wonder whether it's superior. That may depend on which version we consider. For instance, rule utilitarianism, as traditionally conceived, defines right action as action that conforms to a rule the general acceptance of which by humanity would have consequences at least as good for humanity as any alternative rule. But what might be valuable *if* everyone else behaved similarly might not be especially valuable—indeed, could be quite bad—if everyone *actually* behaves quite differently. Driving fifty-five miles per hour might be best if everyone else did as well but not if everyone else is driving seventy-five miles per hour. So it may be a mistake for the indirect consequentialist to identify right action with action conforming to rules with optimal *general* acceptance value. Instead, he might identify right action with action in conformity with personal rules having optimal acceptance value, given the way others will actually behave.[3]

Even if the set of actions on the best motives or rules is different from the set of best actions and the former produces more value overall, this still does not favor indirect consequentialism. After all, the direct consequentialist assesses all sorts of things in addition to actions—including persons, policies, institutions, and, notably, motives and rules. Because she assesses all things according to their comparative value, she should prefer having and acting on the best motives or rules to performing the best actions, just in case these diverge and acting from the best motives is best overall. Indeed, any cases in which acting from the best motives or rules produces suboptimal actions would arguably count as cases of blameless wrongdoing.

If there aren't compelling advantages offered by indirect consequentialism, we might focus on more traditional direct forms of consequentialism. But doesn't the direct consequentialist assessment of actions imply that an agent's deliberations should always be guided by a comparative cost-benefit analysis of the alternatives? Whereas such

Subscriber: King's College London; date: 14 February 2019

consequentialist accounting may well be appropriate in special circumstances, it does not seem generally appropriate. For one thing, conscious attempts to optimize are often counterproductive. Optimizing deliberations are often inefficient when they are costly and time for deliberation is scarce. They are also subject to bias. Interpersonal maximization is often distorted by the agent's sense of his own interest, by his investing his own interests with normative significance out of proportion to their magnitude relative to the interests of others. Similarly, intrapersonal maximization is often distorted by the temporal proximity of benefits or harms, by agents investing near-term benefits and harms with normative significance out of proportion to their actual magnitude. Still more important, certain valuable activities and relations, including avocations and intimate associations, seem incompatible with regular scrutiny of their consequential value. As Bernard Williams suggests, the optimizer who pursues various projects or who provides aid and succor to his loved ones only after concluding that this is the optimal use of his resources, instead of merely consulting his passions or his loyalties, seems to "have one thought too many" (1976).[4]

(p. 388)

There are several concerns here, but many of them rest on the assumption that the direct consequentialist should treat her consequentialism as a decision procedure, always deliberating as an optimizer. But a moral theory can supply a *standard* of right conduct, explaining what makes right acts right, without supplying a decision procedure. In *The Methods of Ethics*, Henry Sidgwick notes the so-called paradox of hedonistic egoism that one often better secures pleasure if one does not consciously aim at it (1966, pp. 48, 135–136). As Sidgwick notes, the paradox, if true, tells us something about how to satisfy the hedonistic egoist standard; it is not an objection to that doctrine. Similarly, he notes that satisfaction of the utilitarian standard may require that agents not always deliberate in explicitly consequentialist terms (1966, p. 413; see Bales, 1971; Brink, 1989, pp. 256–262; Railton, 1984). But if it was always counterproductive to reason with consequentialist principles or if it was best for the truth of consequentialism to be known only to a select philosophical elite, then, as Sidgwick notes, consequentialism would have the status of an "esoteric" morality (1966, pp. 489–490). This result would be worrisome, inasmuch as we expect moral principles to play some role in moral deliberation, especially about perplexing cases, and in moral education. But moral principles can play a significant role in moral deliberation without functioning as decision procedures. In particular, moral principles can *regulate* an agent's conduct in variety of ways without always figuring consciously in her deliberations or motivations. A principle will regulate an agent's conduct, even when she doesn't consult it, if she wouldn't act as she does unless her conduct satisfied the principle or might reasonably be thought to satisfy it. So an agent can act with a variety of motives and by consulting a variety of secondary precepts consistently with her conduct being regulated by a different master principle, provided that she so acts when doing so is clearly permissible (nonblameworthy) according to the principle, and provided that she refuses so to act when doing so would clearly be impermissible (blameworthy) according to the principle. However, if the principle does regulate her behavior, she will consult it when her motives and precepts that normally

track the principle give uncertain or conflicting guidance, and she will periodically step back from her everyday motives and precepts and reassess their compatibility with the principle. In this way, the direct consequentialist can recognize that responsible and admirable agents need not and should not constantly consult the consequentialist principle or always engage in conscious consequentialist analysis, provided that their behavior is suitably regulated by consequentialist principles. If so, worries that consequentialism requires a mindset of moral accountancy that is inconsistent with spontaneity, authenticity, or fidelity appear misplaced or, at least, premature.[5]

(p. 389) 4. Perfectionism and Other Conceptions of the Good

Direct consequentialism assesses actions, motives, persons, policies, and institutions in terms of the good they produce. But consequentialism, as such, does not tell us what is good. For that we need a theory of value. Though some of my discussion abstracts from different evaluative assumptions to focus on consequentialism as such, this agnosticism about the good is not always possible or helpful.

It is an interesting question whether there are any impersonal goods. I am somewhat skeptical that we would recognize anything as valuable independently of any contribution that it makes to improving lives, whether human, rational, or sentient. In any case, my discussion will focus on conceptions of a person's good. It is common to identify a person's good with his interests, well-being, or welfare. We can even identify a person's good with his happiness, provided that we do not assume at the outset that happiness is conceptually tied to satisfaction or contentment.[6] We can think of a person's good as what we ought to care about intrinsically, insofar as we are concerned about him for his own sake. We can then recognize different substantive conceptions of the good for a person (his interests, well-being, welfare, or happiness). Hedonism and preference-satisfaction views are subjectivist conceptions, whereas perfectionism and objective lists are objective conceptions. Conceptions of the good might also be mixed, containing both subjective and objective elements. My working assumption will be that pure subjectivist conceptions of the good are implausible and that some more objective conception in which perfectionist elements play a significant role is most promising. Let me briefly sketch this assumption.[7]

Hedonism is a form of *extreme subjectivism*; it says that happiness or value consists in mental states or sensations alone.[8] The desire-satisfaction theory, by contrast, is a form of *moderate subjectivism*, because it says that happiness depends upon a person's mental states—her desires—but consists in the satisfaction of her desires.

Familiar thought experiments show why it is difficult to maintain, as hedonism requires, the extreme subjectivist claim that happiness or value consists in psychological states alone. Robert Nozick questions whether we would really choose to hook up to an experience machine that provides experiences of any life we would enjoy; he assumes that we want to be the authors of our own lives, make real differences in the world, and

Subscriber: King's College London; date: 14 February 2019

sustain meaningful relations with others, and not merely have experiences as if we were doing these things, no matter how pleasant such experiences might be (1974, p. 42). In Aldous Huxley's *Brave New World* (1946), Deltas and Epsilons form the working classes, who are genetically (p. 390) engineered and psychologically programmed to acquiesce in and indeed embrace intellectually and emotionally limited lives that are liberally seasoned with mood–altering drugs. In such lives, pleasure and contentment are purchased at the price of dignity.

The experience machine raises problems for extreme subjectivism, but is not a direct threat to moderate subjectivism. For the desire–satisfaction theorist can note that the experience machine does not satisfy its clients' desires to be and do certain things (though it does, *ex hypothesi*, leave them satisfied). However, we sometimes judge that people who *are* satisfying their deepest desires nonetheless lead impoverished lives, because their desires are for unimportant or inappropriate things. Deltas and Epsilons lead contented lives precisely because they are satisfying their chief desires. While a certain amount of realism in one's ambitions and desires may be a good thing, we do not (in general) increase the value of our lives by lowering our sights, even if by doing so we increase the frequency of our successes.

Insofar as the goals of Deltas and Epsilons are based on false beliefs about their capacities or are the result of brainwashing, the moderate subjectivist may think that our concerns can be met by appealing, not to actual preferences, but to suitably *idealized* preferences that are fully informed and formed under conditions free from psychological manipulation by others. On this sort of view, what is good for someone is what his idealized self would want his (nonidealized) self to want.[9]

However, laundering people's preferences is an inadequate remedy. An ideal appraiser, like John Stuart Mill's competent judge, is supposed to be fully informed about all aspects of all the possibilities open to her. But there are various questions about the coherence and relevance of fully informed desire. Can one coherently combine wildly disparate possible experiences in one overall evaluative perspective (see Rosati, 1995; Sobel, 1994)? Moreover, one can't rule out the possibility that full confrontation with the facts wouldn't extinguish desire or shape it in ways that one would pretheoretically identify as pathological (see Gibbard, 1990, p. 20). Furthermore, we may wonder whether idealized desire satisfaction views don't confuse what's in our interest and what interests us (see Darwall, 1997). For it's not clear that everything that one might reasonably (or not unreasonably) desire would contribute to one's good.

The idealized desire-satisfaction view also faces a serious dilemma. If the process of idealization is purely formal or content-neutral, then it must remain a brute and contingent psychological fact whether suitably idealized subjects would care about things we are prepared, on reflection, to think valuable. But this is inadequate, inasmuch as we regard intellectually and emotionally rich lives as unconditionally good and intellectually and emotionally shallow lives as unconditionally bad. For a person with the normal range of intellectual, emotional, and physical capacities, it is a very bad thing to lead a simple

and one-dimensional (p. 391) life with no opportunities for intellectual, emotional, and physical challenge or growth. One's life is made worse, not better, if, after informed and ideal deliberation, that is the sort of life to which one aspires.

Alternatively, we might conclude that anyone who would endorse shallow and undemanding lives simply could not count as ideal appraiser. We might agree with Mill, who claims in *Utilitarianism* that any competent judge, who has a proper sense of his own dignity, would never approve of contented but undemanding lives (1978b, chap. 2, para. 6). But if this is to explain how such lives are categorically bad, then it must be that one won't count as an ideal appraiser unless one possesses a sense of dignity that reflects a belief in the value of activities that exercise one's higher capacities. But such a notion of idealization carries substantive evaluative commitments. Suitably idealized desire, understood this way, presupposes, rather than explains, the nature of a person's good.

These worries about extreme and moderate subjectivism lend plausibility to objective conceptions of the good. One form of objectivism is a list of objective goods, such as knowledge, beauty, achievement, friendship, and equality.[10] Some such list may seem the only way to capture the variety of intrinsic goods. But if it is a mere list of goods, with no unifying strands, it begins to look like a disorganized *heap* of goods.[11] While we can't assume that there is a unified account of the good that is reflectively acceptable, we have reason to look for one and treat a mere list of objective goods as a kind of fallback position.

One promising objective conception that goes beyond a mere list of goods is *perfectionist*. There is a venerable perfectionist tradition, common to Aristotle, Mill, and T. H. Green, among others, that identifies a person's good with the perfection of her nature.[12] But human nature can be conceived of as a biological or normative category. It is hard to find capacities that we have as a biological species that are essential and whose exercise seems distinctively valuable.[13] A more promising avenue is to understand the appeal to human nature in normative terms. In *An Essay Concerning Human Understanding*, John Locke distinguishes between persons and men (or, as we might prefer to say, human beings) and claims that the concept of a person is a "forensic" or normative concept (1979, II.xxvii. secs. 8, 15, 17–21, 23, 26). Part of what Locke means is that only persons are accountable in law and morality, because only persons are responsible for their actions. Nonresponsible agents, such as brutes and small children, act on their strongest desires; if they deliberate, it is only about the instrumental means to the satisfaction of their desires. By contrast, responsible agents must be able to distinguish between the intensity and authority of their desires, deliberate about the value or authority of their desires, and regulate their actions in accordance with their deliberations. On this view, what is essential to persons are these capacities for practical deliberation and regulation of the will that mark one as a responsible agent. It is significant that the main figures in the perfectionist tradition understand the essentials of human nature in something like this normative way.

(p. 392)

This kind of perfectionist view claims that a person's good consists in activities that exercise and express her capacities for practical deliberation. Such a view explains why we value lives of various sorts in which people are self-directed and engaged in activities that exercise their creative powers. In doing so, the perfectionist is able to accommodate an attractive kind of *pluralism* about the good, which recognizes a variety of different but equally or incommensurably good lives, without lapsing into an unsustainable *content-neutrality*, which places no substantive restrictions on the content of a good life. This sort of perfectionism also explains our reservations about shallow and undemanding lives, even when these are successful in meeting the agent's actual or reflective aims and aspirations. Finally, this sort of perfectionism is well suited to answer an important question about the normativity of the good. Though many people fail to care about what is actually good for them, it is common to think that people would, or at least should, care about their own good if they understood what it consisted in. If so, we can ask about any putative conception of an agent's good *why* he should care about it. Any account of the good should be able to explain why it is reasonable or makes sense for a person to care about his good, so conceived. It is not obvious why one should aim to experience pleasure or satisfy desires, regardless of the source of the pleasure or the object of desire. By contrast, a perfectionist conception that stresses the exercise of deliberative capacities ties the content of the good to the very capacities that make one a responsible agent, subject to reasons for action, in the first place. In pursuing this sort of perfectionist good, one is exercising the capacities that make one a rational agent. This kind of perfectionism, it seems, promises to explain the normativity of the good.

5. Accommodation and Reform

So far, I have examined various theoretical choices the consequentialist must make, expressing special interest in direct, maximizing conceptions of consequentialism that give an important role to perfectionist goods. However, I have not yet addressed the issues about whose well-being matters, and how it matters, that separate impartial (agent-neutral) and partial (agent-relative) consequentialists. I will examine these forms of consequentialism separately and in some detail.

My test of adequacy will be systematic comparative plausibility. Does the view in question recognize or violate plausible constraints on an adequate moral theory, and does it cohere well with other independently plausible philosophical commitments? In addition to various theoretical virtues, a plausible theory must cohere well with our independent moral judgments about actual and hypothetical (p. 393) cases. A good theory aims to subsume and explain familiar moral precepts, but the theory that does this best and has various theoretical virtues may well be morally revisionary. Ideally, we make tradeoffs among our theories, considered judgments, and other philosophical commitments, making adjustments here and there, as overall coherence seems to require, until our ethical views are in *dialectical equilibrium*.

Such a dialectical examination, therefore, involves both accommodation and reform of our preexisting moral outlook. When examining consequentialist conceptions, I will show that they sometimes appear to have counterintuitive implications. The consequentialist has two main responses available. He can respond by arguing that consequentialism can, after all, accommodate the allegedly recalcitrant intuition. Alternatively, where accommodation is impossible, the consequentialist can urge us to reform our intuitions, either because the intuitions lack an adequate philosophical rationale, or because the demands of global accommodation require local reforms. We will have to decide how much accommodation is possible and how much reform is reflectively acceptable.

6. Utilitarianism and Impartiality

Contemporary discussions of consequentialism almost always focus on impartial or agent-neutral consequentialism, which tells agents to promote the good, as such, and not just the good of the agent or some other limited class of people. If we do not recognize impersonal goods, then impartial consequentialism directs us to what would most advance the well-being or happiness of all affected parties. This is the central claim of utilitarianism, though different conceptions of utilitarianism result from different conceptions of well-being or happiness.

Why take utilitarianism or any other form of impartial consequentialism seriously? Bentham and Moore seem to have thought that it is an analytic truth that one ought to do the action with the best consequences (Bentham, 1970, ch. 1, sec. 11; Moore, 1903, secs. 17, 89). But neither agent-neutral consequentialism nor utilitarianism passes Moore's own test for analytic truths—the Open Question Argument—because it is possible for competent speakers to doubt whether right action and action that maximizes value or happiness are the same. Others have thought that utilitarianism is attractive because it recognizes the central importance of benevolence as a virtue and the important role of sympathy in moral motivation (see Boyd, 1989, pp. 215–216; Scanlon, 1982, p. 115). But few think that benevolence is the only or the most important virtue.

My own view is that the chief attraction of utilitarianism lies in its interpre (p. 394) tation of the concept of *impartiality*. It is a salient feature of modern conceptions of morality that they aim to overcome parochial concern. It is common to think of the moral point of view as one that asks an agent to transcend his own private concerns and allegiances. We might understand such transcendence in terms of adoption of a point of view that is impartial as among the interests of affected parties. The utilitarian conception of impartiality says that each is to count for one and none for more than one. When utilitarianism was first championed in the eighteenth and nineteenth centuries, its impartiality made it part of a revisionary moral and political theory that tended to undermine familiar institutions of class and privilege. This moral reform is now generally thought to have been a progressive influence, correcting an indefensible moral parochialism.

Some Forms and Limits of Consequentialism

The utilitarian conception of impartiality assigns no moral importance, as such, to whom a benefit or burden befalls; it is the magnitude of the benefit or harm that matters morally. This conception of impartiality supports a maximizing moral standard. The utilitarian takes everyone's interests into account by aggregating their interests, balancing benefits to some against harm to others, as necessary, so as to produce the best total outcome.

Some critics object to the utilitarian conception of impartiality as requiring interpersonal balancing. Whereas balancing goods and harms may be acceptable *within a life*, many think that it is not acceptable to balance goods and harms *across lives*. On the aggregative conception, individual claims may simply be outvoted by a majority. In order to respect the separateness of persons, critics claim, our concern for each person must take a distributed, rather than an aggregative, form. One such distributed conception of impartiality is *contractualism*, which claims that distributions of benefits and harms must be acceptable, in the relevant sense, to *each*. One version of contractualism claims that actions and the way they distribute benefits and harms are right insofar as they conform to principles that no one can reasonably reject (Scanlon, 1998, p. 153). By giving each person in effect a veto, the contractualist seeks a kind of unanimity, in contrast to the majoritarianism of utilitarianism. The interpersonally best option may usually be acceptable to many, but it can fail to be acceptable to each.

How best to model impartiality is a large and important topic that goes beyond the scope of this study. But it is worth noting that utilitarian and contractualist conceptions of impartiality need not be treated as mutually exclusive alternatives. Given that people's actual talents, holdings, and prospects are often the product of arbitrary forces within natural and social lotteries, for which the individual has little responsibility, it would often be unfair in the distribution of benefits and burdens to give everyone a veto on the basis of his actual position and preferences. One needs to *moralize* the contract. One needs to replace the question "What arrangements could no one reject given knowledge of his actual endowments and preferences?" with something like the question "What arrange (p. 395) ments could no one reject in fair circumstances that abstract from morally arbitrary facts about his endowments and preferences?" This arguably requires replacing the idea of an ex post agreement among different individuals with conflicting interests with an ex ante choice of a single self-interested individual under a veil of ignorance about his actual endowments and preferences. If we model contractualism this way, it is arguable that contractors would choose so as to maximize expected total or average welfare, for such a principle, in contrast with nonmajoritarian principles, is antecedently more likely to advance one's interests once the veil is lifted.[14] If so, contractualists need not reject the utilitarian conception of impartiality.

However, to say that the utilitarian conception of impartiality is compatible with contractualist conceptions is not to say that the utilitarian conception is unproblematic. We can group together several concerns about utilitarian impartiality under two main headings—*constraints* and *options* (see Kagan, 1989, ch. 1). Constraints are moral prohibitions, which are often thought to correlate with moral entitlements that

individuals possess—such as rights—that limit what someone may do to them, even in the pursuit of good consequences. On such views, it can be wrong to do something, even though doing so might maximize value. Other critics have focused on options, rather than constraints, alleging not that the consequentialist demand for the agent to promote the good violates duties to others, but rather that it ignores prerogatives that the agent has to devote attention and resources to her own projects and those of others with whom she is associated out of proportion to their impersonal value.

7. Impersonal Constraints

One source of concern about utilitarian impartiality is its apparent failure to accommodate *impersonal* constraints. These are duties that an agent owes to anyone regardless of the relationship in which she stands to that person. Typically, these duties are correlated with claims or rights that a person has to be treated or not to be treated in certain ways. On one understanding, these duties are not to be violated, even if doing so produces more value overall. Nozick emphasizes this aspect of rights, when he insists that rights be understood as side constraints rather than as important goals; it is wrong for an agent to violate one person's right, even if so doing would minimize the total number of violations of such rights by others (1974, p. 29). To treat rights as side constraints is to recognize values that should be honored rather than promoted.

(p. 396)

Most everyone recognizes rights, and the conception of rights as side constraints is quite appealing. But, on reflection, side constraints can appear paradoxical. As Nozick himself notes, if the nonviolation of a constraint is so important, shouldn't we take as our goal the minimization of violations of that constraint (1974, pp. 30–31)?

Nozick's own answer is to appeal to the separateness of persons and the Kantian demand that we treat all agents as *ends* and never merely as means. But the Kantian requirement does not obviously require side constraints. Suppose that only by causing harm to B can A prevent individually comparable harms to C, D, and E. If A harms B only in order to protect C, D, and E, perhaps A treats B as a means, but he need not treat her as a mere means. To do that would require viewing her as a mere instrument or tool, not as someone whose own agency is valuable. But A need not view her that way. He can take her agency into account; if so, he proceeds, but with reluctance that derives from a concern with her agency. If A could have protected C, D, and E without harming B, he certainly would have. If A acts impermissibly in acting so as to minimize harm, it is not because in so acting he must be treating those whom he harms as mere means.

Sometimes friends of side constraints appeal to a sort of *inviolability* that individuals possess if and only if their fundamental interests are protected by side constraints (Kamm, 1996, vol. 2, pp. 271–278; Nozick, 1974, p. 31; Rawls, 1971, pp. 3–4). But to make B inviolable in this way will require turning a deaf ear to the comparable interests of C, D, and E. This seems to deny them moral *considerability*. Though we want to take seriously

the fundamental interests of each, it is not obvious that we should endorse inviolability, because ensuring the inviolability of each denies the moral considerability of others.[15]

Furthermore, we may wonder whether impersonal constraints would be acceptable within a suitably moralized contract. If we appeal to an ex ante self-interested choice subject to ignorance about whether one will be A, B, C, D, or E when the veil of ignorance is lifted, then there is every reason to believe that one would prefer a harm–minimization principle to one representing a side constraint upon causing harm. For, all else being equal, one clearly stands a better chance of avoiding harm under harm minimization than under a side constraint. If so, then impersonal constraints may seem problematic from the point of view of contractualist impartiality, as well as utilitarian impartiality.

So, despite the initial intuitive plausibility of impersonal constraints, they are not unproblematic. Absent an attractive rationale for such constraints, it would be premature to reject utilitarianism for its failure to accommodate them (see Kagan, 1989; Scheffler, 1982, ch. 4).

(p. 397) 8. Personal Options

Whereas some critics of utilitarianism focus on constraints, others focus on options. Utilitarian impartiality demands that an agent always act so as to bring about the impersonally best outcome. But especially when we recognize the variety of grave imperfections in the world and the opportunities that these imperfections provide for contributing to a better world, utilitarianism can seem very demanding indeed. So much so that we may begin to wonder whether utilitarianism leaves the agent room to pursue those projects and associations that she cares most about and that give her life meaning. Williams has brought to our attention worries of this sort about the conflict between impartial moral conceptions, such as utilitarianism, and the personal point of view (1973, 1976). Responding to this conflict, Samuel Scheffler has proposed to *moderate* the demands of utilitarianism by recognizing moral options or *prerogatives* on the agent's part to devote time, energy, and resources to her own projects out of proportion to their impersonal value (1982, esp. chs. 1–3, 1992, esp. chs. 6–7).

Recognition of personal options is one way to recognize a limit on the sacrifices that morality can demand. And, unlike impersonal constraints, Scheffler argues, options are not inherently paradoxical. An important rationale for options is that they allow the agent to integrate morality into a reasonable life plan. In order for moral demands to be integrated into a reasonable and satisfying life plan, they must be motivationally accessible to agents. But, Scheffler argues, the "natural independence of the agent's point of view" means that agents have concerns for themselves, their own projects, and their intimates that is out of proportion to their impersonal value. But then impartiality without options won't be motivationally accessible to agents; only a form of impartiality moderated by options can be integrated into a reasonable and satisfying life plan.

One interesting question is whether one can have options without some sort of constraints, as Scheffler proposes. Scheffler seems to think of options or permissions as shielding the agent from the demands of impartial consequentialism. But if they are to provide a significant shield, it seems that an agent's options should correlate with duties of others to respect the agent's nonoptimizing personal choices (see Alexander, 1987). It is true we can imagine a system of permissions without any correlative duties in which each is free to pursue his own personal projects but no one is under a duty to refrain from interfering with the projects of others. But such unsupported options are not very attractive, especially to someone who saw options as a way of protecting the agent from demands by others. If so, then it is hard to defend options without constraints. Whether such constraints have to be understood as side constraints is, of course, a separate matter.

Should we recognize personal options? While I have no impossibility proof (p. 398) to offer, the arguments for them are not obviously compelling. It is not clear that utilitarianism fails the test of motivational accessibility. That may depend on how we understand the test. If motivational accessibility is relativized to people's actual motivations, then it may well be true that utilitarian demands are motivationally inaccessible to many, inasmuch as many, no doubt, do care about their own projects and commitments out of proportion to their impersonal value. But motivational accessibility, so understood, has potentially conservative implications, severely limiting the demands of moral reform. If this is how we interpret motivational accessibility, we may well decide to reject it as an acceptable constraint on moral theory. Alternatively, motivational accessibility might be relativized to possible or desirable motivations. But then a utilitarian morality may not be motivationally inaccessible. Motivation can be responsive to moral and other normative beliefs. But then if there are good arguments for an impartial morality, such as utilitarianism, acceptance of these arguments can help produce motivation congruent with such demands.[16] In short, it is hard to identify a conception of the motivational accessibility requirement that both yields a plausible requirement and clearly rules out utilitarian conceptions of impartiality.

9. Personal Constraints: Associative Duties

In some ways, what is most puzzling is the thought that the personal point of view limits the demands of impartiality by way of options. Insofar as commonsense morality recognizes limits on impartial demands, it recognizes duties, and not just permissions, of a personal nature. I am under duties of self-cultivation and duties toward associates that limit the impersonal good I can be expected to promote. I have in mind what are sometimes called *special obligations* that an agent has toward himself and toward others to whom he stands in various sorts of special relationships. Different kinds of special obligations are rooted in different sorts of relationships or associations—including parent-child relationships, marriage, friendships, and professional relations. Some of these relations are undertaken in a wholly voluntary way (as when I choose a spouse, friends, or colleagues), whereas others appear to have significant nonvoluntary aspects (I am

unable to choose my parents). No doubt the nature and texture of such associations are quite variable. Nonetheless, there are common themes of shared experiences, learning from another, mutual trust, cooperation, common aims, and mutual concern pervading such associations. We might characterize associations (p. 399) as involving shared history between people that obtains when the beliefs, desires, intentions, experiences, emotions, and actions of each interact with and influence those of the other. Indeed, it would be natural to think that the strength of an association is proportional to the degree of psychological interaction and interdependence, with stronger and more intimate associations held together by greater psychological interdependence and influence. One might think of one's associational relations as forming a set of concentric circles in which my closer associates lie on the inner circles and more remote associates lie on the outer circles. But if special obligations are based on associational ties, then it would be natural for the strength of associational duties to be proportional to the strength of the underlying associational bonds.

How should I express concern for myself and my associates? That depends on which theory of the good is correct. If, as I have argued, perfectionist ingredients form the central elements of a person's good, then I should express my concern for my associates by doing things to further the proper development of their deliberative competence and their pursuit of projects and plans that they have reflectively endorsed and that exercise their deliberative capacities.

If we understand associational duties on this model, then such duties depend upon the right sort of interpersonal interaction and influence and do not automatically arise from all interpersonal relations. So, for example, children would owe no typical filial duties to biological parents who have played no role in their nurture and development. Similarly, estranged spouses would not have typical marital obligations toward each other. Hermits who live in physical proximity to each other would not be obligated as neighbors. Insofar as these restrictions on the scope of associative duties are reflectively acceptable, this makes the proposal to ground such duties in interpersonal interaction and influence more attractive.

On this view, associational relations ground special concern for the well-being of one's associates. Acting on this concern will often require modifying the roles that associates play in an association. This conception of associative duties contrasts with some strands within the communitarian tradition that find the content, as well as the ground, of associative duties in past associational relations and imply that associates have a duty to conform to the roles established by past association.[17] Past association may ground a duty of concern, but it does not settle the form that such concern should take. If our past association has not been mutually beneficial, then our shared history gives us special reasons to modify the terms of our relationship so as to be better adapted to the needs of one or both parties.[18] In the limiting case, special concern for the good of associates can provide reasons to discontinue an association, if that is what is best for associates.

Some Forms and Limits of Consequentialism

In his article "Self and Others," Broad describes a moral theory that recognizes associative duties that he calls *self-referential altruism* and associates with common-sense morality. Like utilitarianism, it recognizes a reason to be concerned about anyone whom it is in the agent's power to affect for better or worse, (p. 400) but it insists that the weight or strength of the agent's obligations is a function of the relationship in which she stands to potential beneficiaries (1971, p. 280). Perhaps associational bonds also create options, but, as Broad recognizes, they characteristically generate obligations or duties. We have duties toward associates to enable and assist their development, to facilitate their projects and plans, to protect them from certain dangers, to console them in times of need, to provide constructive criticism, and so on. All else being equal, our duties toward associates take precedence over duties to nonassociates, and our duties to closer associates take precedence over duties to more remote associates. This aspect of self-referential altruism is hard to square with utilitarianism. The problem is that special obligations involve duties to associates whose normative strength appears to be out of proportion to the impersonal good that their fulfillment embodies.

Consider Sidgwick's admirably clear-headed attempt to accommodate special obligations within his hedonistic utilitarian framework. In the *Methods of Ethics*, he argues that the recognition of special obligations and a differentially greater concern for those to whom one stands in special relationships is in general optimal, because we derive more pleasure from interactions with associates, we often have better knowledge about how to benefit associates, and we are often better situated causally to confer benefits on associates (1966, pp. 431–439).

However, even if the demands of special concern and impartial concern often coincide, the coincidence is imperfect. I may derive more pleasure from interaction with my associates than with strangers, but those who are strangers to me have their own associates who derive special pleasure from them. If so, it is not clear how an impartial concern with happiness explains why I would have any reason to privilege the claims of *my* associates over those who are strangers to me but associates of others. Moreover, often—where the beneficiaries are near at hand and the benefits in question are fairly obvious—I am just as well positioned epistemically and causally to benefit strangers as to benefit my associates. When this is so, the classical utilitarian has no reason to regard an agent's investments in his friends as a more efficient use of his resources.

These accounts of special concern within an impartial or impersonal perspective appear to be unable to give a sufficiently robust account of special concern. The problem is that utilitarianism's impartiality assigns only *extrinsic* significance to special concern; special concern is valuable only so far as it tends causally to promote human happiness. By contrast, common sense attaches *intrinsic* significance to special relationships; the fact that A and B are friends gives A special reason to be concerned about B that he does not have to be concerned about C.[19]

Alternatively, we might put this point in terms of the distinction between agent-relative and agent-neutral reasons. Reasons to promote the good, as such, are agent-neutral

reasons, whereas reasons to promote the good of those to whom the agent stands in special relationships are agent-relative reasons. Association (p. 401) seems normatively significant, because it seems to transform the reasons the agent has independently of the association. If so, one's reasons to be concerned about one's associates are agent-relative, not agent-neutral.

Because associational duties assign intrinsic and agent-relative significance to the shared history among associates, they resist capture within a utilitarian conception of impartiality. Special concern may not always trump impartial demands to promote happiness; but the former cannot be reduced to the latter. If so, we might entertain utilitarian or consequentialist views as revisionary challenges to the legitimacy of special concern but not, I think, as justifications of special concern.

10. Voluntarist and Distributive Concerns about Associative Duties

Many would think that this is reason to conclude that however adequate utilitarianism is as an account of impartiality, it represents an inadequate account of the sort of partiality characteristic of associational duties. However, while conceding the intuitive appeal of associative duties, some complain that such duties, like impersonal constraints, are paradoxical and require an adequate philosophical rationale that explains their normative significance. In particular, Scheffler has argued that associative duties are problematic on two fronts. Associative duties appear to be overly demanding of agents when, as in some familial relationships, they obligate agents to have special concern for associates they have not sought out. Recognizing such duties appears to violate the *voluntarist* assumption that all duties must be voluntarily undertaken by the agent. Whereas the voluntarist is worried about the *costs* of association for the agent, there is *distributive* concern about the *benefits* of association. Precisely insofar as associative duties give more urgency to the claims associates make on each other, they reduce the comparative urgency of the claims of nonassociates on associates. Associative duties privilege the claims of insiders against those of outsiders, and so might seem to give rise to legitimate complaints by outsiders. These objections render associative duties problematic and in need of an adequate philosophical rationale.[20]

But it is hard to see these as decisive objections to associative duties. First, we might not be as concerned by either the costs or the benefits of associative duties if we bear in mind that such duties involve *both* costs and benefits. Insider privileges may seem less significant when they are balanced against insider burdens, and insider burdens may seem less onerous when insider benefits are reck (p. 402) oned in. Moreover, it is not clear that associative duties, as understood here, violate the voluntarist assumption. Because associative duties, on this view, do not arise from just any interpersonal relations but require interpersonal interaction and influence (see section 9), they depend upon the voluntary actions of associates and so cannot be wholly nonvoluntary. Furthermore, the voluntarist assumption that duties can be generated only by the agent's voluntary

undertakings itself stands in need of a rationale. Indeed, voluntarism is flatly inconsistent with utilitarianism and any other moral theory that recognizes various noncontractual duties toward others. So the utilitarian is in no position to complain that associative duties violate voluntarism. The distributive objection does focus on a way in which associative duties require a deviation from egalitarian or impartial concerns, but, of course, this is just the direct consequence of recognizing the demands of partiality. So I doubt that the associative duties that resist capture within the net of utilitarian impartiality are any more problematic than utilitarianism itself.[21]

11. Personal Identity, Utilitarianism, and Associative Duties

One way of defending utilitarianism against worries about constraints actually appeals to some of our claims about associative relations together with claims about personal identity. Several worries about utilitarianism's liability to violate rights focus on the *person neutrality* of utilitarianism. The utilitarian conception of impartiality assigns no moral importance, as such, to whom a benefit or burden befalls; it is the magnitude of the benefit or harm that matters morally. The utilitarian takes everyone's interests into account by aggregating their interests, balancing benefits to some against harm to others, as necessary, so as to produce the best total outcome. If the magnitude of benefits and harms is of moral importance as such, but their distribution across lives is not, then one should maximize net value, rather than seek any particular interpersonal distribution.

As such, person neutrality effects a kind of impartiality across lives akin to the impartiality that *temporal neutrality* effects within lives. It is a common view that the temporal location of a benefit or harm within a life should not, as such, have any rational significance. A person should only be concerned with the magnitude of the benefit or harm within her life, not its temporal location, which implies that she should be impartial among different stages of her own life and maximize her overall good, rather than achieve any particular intertemporal distribution.

(p. 403)

Indeed, many have seen the motivation for utilitarianism as extending the familiar balancing and maximizing procedure from diachronic intrapersonal contexts into interpersonal contexts (Rawls, 1971, pp. 23–24; Sidgwick, 1966, p. 382). But, as I showed earlier (section 6), some critics of utilitarianism object to this assimilation of interpersonal balancing to intrapersonal balancing. Whereas balancing goods and harms may be acceptable *within a life*, many think that it is not acceptable to balance goods and harms *across lives*. To engage in interpersonal balancing, as utilitarianism does, is to fail to respect the *separateness of persons* (Nagel, 1970, pp. 134, 138–42; Nozick, 1974, pp. 31–34; Rawls, 1971, pp. 23–29, 187–188; Williams, 1976, p. 3).

This asymmetry between intrapersonal and interpersonal balancing is linked to concerns about *compensation*. In the case of intrapersonal balancing, the sacrifice of one's present good for one's later, greater good is compensated; benefactor and beneficiary are the same. But in the case of interpersonal balancing, benefactor and beneficiary are different people; unless the beneficiary reciprocates in some way, the benefactor's sacrifice will go uncompensated. Whereas intrapersonal balancing is automatically compensated, interpersonal balancing is not. This may make person neutrality problematic in a way that temporal neutrality is not.

If the compensation principle is interpreted so as to forbid all uncompensated sacrifices and all interpersonal balancing, then it apparently forbids all redistributions of resources from the superrich to the destitute—no matter how small a burden on the superrich and how great a benefit to the destitute. So interpreted, the compensation principle is hard to accept.

Derek Parfit has tried to defend utilitarianism against the separateness of persons objection in a different way. He has argued that if we accept a traditional conception of personal identity that analyzes personal identity into psychological relations, then we should deny that the separateness of persons is fundamental. If we reject the separateness of persons, then we can defend the utilitarian conception of interpersonal balancing (Parfit, 1984, ch. 15).

Parfit's view is similar to other views in the Lockean tradition of thinking about personal identity (Locke, 1979, II.xxvii), such as the views of Shoemaker (1963, 1984), Wiggins (1967), and Nozick (1980, ch. 1). Parfit calls his view a form of psychological reductionism. The psychological reductionist analyzes personal identity into relations of psychological *continuity* and *connectedness*. Roughly, two persons are psychologically connected insofar as the intentional states and actions of one influence the intentional states and actions of the other. Examples of intrapersonal psychological connections would include A's earlier decision to vote democratic and her subsequent casting her ballot for the democratic candidate, A's later memories of a disturbing childhood incident and her earlier childhood experiences, and A's later career change and her earlier reevaluation of her priorities. Two persons are psychologically continuous insofar as they are links in a chain or series of people in which contiguous links in the chain are psychologically (p. 404) connected. Both connectedness and continuity can be matters of degree. According to Parfit's form of psychological reductionism, it is the holding of many such relations of connectedness and continuity that unify the different stages in a single life. More specifically, on this view, personal identity consists in maximal (nonbranching) psychological continuity.[22]

But psychological connectedness and continuity are one-many relations; there can be interpersonal, as well as intrapersonal, psychological connections and continuity. Though I am normally most strongly continuous with myself in the future, I can be psychologically continuous with others with whom I interact psychologically. Interpersonal, as well as intrapersonal, psychological continuity is quite common and can be found, to varying

degrees, in all associative relations. Associates interact and help shape each other's mental life; in such relationships, the experiences, beliefs, desires, ideals, and actions of each depend in significant part upon those of the other.

According to psychological reductionism, what normally distinguishes intrapersonal continuity and interpersonal continuity is the *degree* of continuity.[23] There are more numerous and more direct psychological connections—between actions and intentions and among beliefs, desires, and values—in the intrapersonal case than in normal interpersonal cases. And where the connections among links in a chain are all weaker, continuity between any points in the chain will also be weaker. Different interpersonal associations exhibit different degrees of psychological connectedness and continuity.

Insofar as the difference between intrapersonal and interpersonal relations is a difference of degree, not kind, the separateness or diversity of persons is less fundamental than it would otherwise be. But if the separateness of person is not fundamental, and the same sort of glue that unifies a single life can be found, to a lesser degree, holding together different lives, then the asymmetry between intrapersonal balancing and interpersonal balancing may seem to disappear. If so, utilitarianism's interpersonal balancing may be no more objectionable than the sort of intrapersonal balancing demanded by temporal neutrality.

But while psychological reductionism may show that the separateness or diversity of persons is not fundamental, it does not vindicate utilitarianism. This is because there are real differences in degree of continuity and connectedness in the relations which a person bears to herself and others. We can think of the degrees of connectedness and continuity in terms of a set of concentric circles in which the person occupies the inner circle and her various associates stand in outer circles, depending on the strength and number of psychological interactions and interdependence she has with them. If one's reasons for concern track the degree of psychological interaction and interdependence, then, all else being equal, one has more reason to be concerned about closer associates than more distant associates or nonassociates. But this sort of *interpersonal discount rate* is incompatible with utilitarianism's person neutrality; it requires an agent-relative ethical (p. 405) theory that recognizes associative duties very similar in structure to self-referential altruism (see Jeske, 1993).

12. A Rationale for Associative Duties

I can take this case for associative duties one step further by showing how reasonable assumptions about persons and personal identity provide a normative rationale for associative duties. We saw that Locke claimed that the concept of a person is a normative or "forensic" concept (section 4). In both morality and law, persons are responsible agents; it is only persons who are properly praised and blamed, because it is only persons who have the requisite capacities for practical deliberation. On this view, personhood requires responsibility, which requires deliberative capacities, which require a conception of oneself as a temporally extended self endowed with deliberative capacities. If we view

persons from this Lockean perspective, then it is natural to endorse a version of psychological reductionism that understands the persistence of persons in terms that emphasize the continuous employment of deliberative faculties in the regulation of thought and action. If persons are essentially responsible agents, then an essential ingredient in psychological connectedness must be *deliberative* connections that hold among actions, intentions, and prior deliberations in the deliberate maintenance and modification of intentional states and in the performance of actions that reflect these prior deliberations.

While we normally find maximal continuity within single lives, we have seen that there are significant forms of continuity *across* individual lives within friendship and other forms of interpersonal association. If so, one's relations to associates are similar in kind to, if different in degree from, those that hold between oneself now and oneself in the future. But this suggests that one has the same sort of reasons to be concerned about associates as one does about one's own future self.

How does this help provide a rationale for associative duties? For one thing, it suggests that associative duties are no more problematic than the demands of prudence. Most people, even many utilitarians, recognize the requirements of prudence as normatively significant. But then associative duties have as strong a rationale as the demands of prudence do. We can go further. For we can ask about prudence, as we can about any putative normative standard, why we should care about its dictates. Why should I care about promoting my own good? Appeal (p. 406) to a deliberative conception of the person helps explain how prudence, conceived in perfectionist terms as exercising one's deliberative capacities, is a requirement of practical reason. For when prudence is understood in deliberative terms, it aims at the exercise of the very deliberative capacities that make one a responsible agent in the first place, capable of having and acting on reasons for action. Moreover, if my persistence depends upon the extension of my deliberative control into the future, we can see how the exercise of my deliberative capacities is part of my welfare. But insofar as associational relations involve interpersonal analogues of these psychological and deliberative connections, the resulting reasons to be concerned about the interests of associates will likewise depend upon the very deliberative capacities that make me a responsible agent, subject to reasons for action. On this view, associational bonds manifest the very same psychological relations that make one an agent, and this explains one's reasons to be concerned about associates. If so, we have the makings of a satisfying philosophical rationale for the normative significance of associational bonds. Whether there is any comparably satisfying rationale for utilitarianism remains to be seen.

13. Associative Duties and the Limits of Impartiality

Utilitarianism is plausible insofar as it provides a natural interpretation of the modern ideal of impartiality. But utilitarianism is doubly impartial. Christian and Enlightenment moral conceptions are impartial in the sense of insisting on the *wide scope* of moral concern. The utilitarian conception of impartiality embodies wide scope, because it insists that the scope of moral concern should be universal, extending to all human or rational (or sentient) beings. But utilitarianism is impartial in the further sense that it assigns *equal weight* to everyone's good. And it insists on this equal weighting of everyone's interests not just in special contexts—for instance, demanding that governments weigh the interests of their citizens equally—but of all agents in all contexts. The utilitarian can justify deviations from this second sort of impartiality only on pragmatic grounds as an effective strategy for actually better meeting the demands of equal concern.

The arguments on behalf of associative duties challenge any moral conception, such as utilitarianism, that is impartial in both senses. Associative duties show the need for a moral conception that embodies an agent-relative form of partiality. But to reject impartiality altogether would be to throw out the baby with the bath water. The interesting question is whether we can articulate a moral conception (p. 407) that combines wide-scope impartiality with the sort of partiality embodied in associative duties.

One attractive way to combine impartiality and partiality is self-referential altruism. Self-referential altruism is impartial and agent-neutral insofar as it insists that an agent has nonderivative reason to benefit anyone whom it is within her power to help. But it is partial and agent-relative insofar as it insists that the weight or strength of the agent's obligations is a function of the relationship in which she stands to potential beneficiaries. As Broad recognized, self-referential altruism combines impartiality and partiality in a way that resonates with common-sense morality.

But while self-referential altruism combines impartiality and partiality in an intuitively attractive way, it faces a problem about the normativity of its impartial demands. While the authority of any normative standard can be questioned, the question arises in an acute form for any impartial standard that has the potential to ask agents to make significant personal sacrifices to benefit others with whom they are not directly associated. Though self-referential altruism gives priority, other things being equal, to the claims of those to whom an agent stands in special relationships, it recognizes the claims of anyone, regardless of the relationship in which he stands to the agent. If my sacrifice can do enough good for strangers, then self-referential altruism may well claim that it is my duty to make such a sacrifice. It is not uncommon to interpret morality as requiring uncompensated sacrifices. This may just seem to be another aspect of the sort of transcendence of parochialism characteristic of modern moral conceptions. But we may wonder why we should regard such sacrifices as reasonable.

14. Egoistic Consequentialism

If a fundamental and underived commitment to impartiality raises difficult questions about the normativity of impartial requirements, one might explore the possibility of deriving impartial requirements within a fully agent-relative framework. In providing a rationale for associative duties, I showed how the normativity of associative duties might be grounded in interpersonal relations that extend an agent's rational capacities and, on perfectionist conceptions, contribute to the agent's good. That account of the normativity of associative duties was ultimately egoistic. The obvious challenge here is whether a sufficiently robust commitment to impartiality can be justified on egoist foundations.

Ethical egoism claims that it is an agent's moral obligation to do what promotes his own good or welfare. Such a view makes the agent's own good primary, (p. 408) defining other moral notions in terms of it. It represents an agent-relative form of consequentialism. As with agent-neutral forms, this sort of agent-relative consequentialism admits of different conceptions, depending on whether it takes maximizing or satisficing forms and on whether it takes a direct or indirect form. As before, we can focus on the more traditional direct, maximizing conceptions, inasmuch as the objections to such conceptions do not seem compelling and they seem to have resources to make the same distinctions and claims as their rivals. What this focus leaves open is the conception of the good. Much will depend on the conception of the good on which the egoist draws. As I will argue, perfectionist conceptions have the greatest resources for explaining the normative authority of prudential concern and for justifying other-regarding concern with wide scope. To appreciate these claims better, it will be helpful to look at the limitations in more familiar subjective forms of egoism.

15. Strategic Egoism

What account can we give of the morality of other-regarding concern within an egoist framework, if we employ hedonist or preference-satisfaction assumptions about the good? Of course, most of us have significant concerns for the well-being of associates and more generalized sympathies for other members of humanity that structure our desires and condition what we take pleasure in. Insofar as such other-regarding attitudes are strong and widespread, they provide the basis for an egoist justification of other-regarding conduct.

But this egoist justification of other-regarding action appeals to other-regarding attitudes without grounding them; as a result, it seems unable to explain why those who lack these attitudes should cultivate them or why those who do have them should maintain them. We need a more robust and counterfactually stable justification of other-regarding conduct and concern.

A traditional egoist defense of impartiality tries to argue that even those with more self-confined concerns have reason to broaden their concerns, because the demands of other-

Subscriber: King's College London; date: 14 February 2019

regarding morality and enlightened self-interest coincide. The main lines of this story are familiar enough. Much of impartial other-regarding morality involves norms of cooperation (e.g., fidelity and fair play), forbearance, and aid. Each individual has an interest in the fruits of interaction conducted according to these norms. Though it might be desirable to reap the benefits of other people's compliance with norms of forbearance and cooperation without incurring the burdens of one's own, the opportunities to do this are infrequent. Noncompliance is generally detectable, and others won't be forbearing and co (p. 409) operative toward those who are known to be noncompliant. For this reason, compliance is typically necessary to enjoy the benefits of others' continued compliance. Moreover, because each has an interest in the cooperation and restraint of others, communities will tend to reinforce compliant behavior and discourage noncompliant behavior, with the result that well-socialized individuals will have internalized these norms. If so, compliance is often necessary to avoid the costs of external and internal sanctions. Whereas noncompliance may secure short-term benefits that compliance does not, compliance typically secures greater long-term benefits than noncompliance. In this way, compliance with other-regarding norms of cooperation, forbearance, and aid might be claimed to further the agent's interests. Insofar as this is true, the egoist can ground other-regarding sentiments and explain why those who do not have them should cultivate them and those who do have them should maintain them.[24]

The main problems with this strategic justification of other-regarding conduct and concern involve its scope and stability. The strategic egoist can justify other-regarding duties only toward partners in systems of mutual advantage. But it is a common modern view that morality has wide scope; it imposes obligations of restraint and aid where the agent stands to gain nothing strategically from the cooperation or restraint of the beneficiary. So, for instance, on this view a person can apparently have no reason to be concerned about future generations. And if the wealthy and talented have sufficient strength and resources so as to gain nothing by participating with the weak and handicapped in a system of mutual cooperation and forbearance, the former can have no reason, however modest, to assist the latter. When morality itself is interpreted in terms of strategic egoism, these are counterintuitive limitations in the scope of moral demands.

Moreover, serious limitations remain in the scope and stability of the concern that the strategic egoist can justify toward strategic partners. Sometimes noncompliance would go undetected; and even where noncompliance is detected, the benefits of noncompliance sometimes outweigh the costs of being excluded from future cooperative interaction. Furthermore, even if the coincidence between morality and self-interest were extensionally adequate, it would be counterfactually fragile. On this justification of compliance with other-regarding norms, compliance involves costs, as well as benefits. As a result, it must remain a second-best option, behind undetected noncompliance, in which one enjoys the benefits of others' compliance without the costs of one's own. So, as Glaucon and Adeimantus point out in Plato's *Republic*, if only I was able to enjoy the benefits of the compliance of others without the costs of my own compliance, then I would have no reason to comply (359b8–360d8). But moral norms seem counterfactually stable—they would continue to apply in these counterfactual circumstances—as other-

regarding norms that the strategic egoist can justify are not. This counterfactual instability represents a further limitation in strategic egoism.

So, despite the promise of strategic egoism to justify impartial concern on (p. 410) agent-relative foundations, strategic egoism is unable to justify other-regarding demands with wide scope or stable significance. Such a view purchases normativity, if at all, at the price of failing to recognize impartiality and individual transcendence. Indeed, insofar as one may question the normative authority of pursuing pleasure or satisfying desire, independently of the sources of one's pleasure or the content of one's desires, one may doubt that strategic egoism can claim normative authority.

16. Perfectionist Egoism

We saw that a suitable perfectionist conception of the good promises to explain how prudence, so conceived, has normative authority (section 12). When prudence is understood to aim at the perfection of the agent's deliberative capacities, it aims at the exercise of the very capacities that make one a rational agent in the first place. If such a perfectionist conception of prudence has normative authority, it is worth asking how far such a perfectionist conception of egoism can go in accounting for impartial moral demands.

In addressing this issue, I should reiterate the most important of my earlier claims about persons, associative relations, and the reasons that persons have to care about their associates. If I endorse a version of psychological reductionism, then psychological interaction and interdependence is the glue that unifies disparate stages in a single life. Future directed self-concern involves being concerned about future selves that are uniquely continuous with my present and past selves. However, we saw that such uniqueness is really just a matter of degree of continuity and not a deep fact. Interpersonal, as well as intrapersonal, psychological continuity is not only possible but common. It is found, to different degrees, in all kinds of interpersonal associations in which the intentional states and actions of associates influence each other. Just as an agent should regard the good of his future self as part of his overall good, so too, I argued, should he regard the good of his associates as part of his overall good. This provides a prudential or egoist justification for an agent to be concerned about his associates.

Before examining the nature or scope of such other-regarding concern, we need to better understand the prudential value of interpersonal association. Once I have associates, my reasons to be concerned about them are, on this view, the same sort of reasons that I have to be concerned about my own future good. So perhaps I can see how a concern for my own good requires a concern for their good. But surely the nature and extent of my associations with others are matters (p. 411) that are, at least to some extent, within my control. What prudential reason do I have to cultivate associations with others in the first place?

The perfectionist answer must be that interpersonal association of the right sort makes for the fuller realization of my deliberative capacities. Though I am essentially a deliberative agent, I am not self-sufficient but am cognitively limited. Interpersonal association helps me transcend these limitations. Sharing thought and discussion with another diversifies my experiences by providing me with additional perspectives on the world. By enlarging my perspective, it gives me a more objective picture of the world, its possibilities, and my place in it. This echoes Plato's and Aristotle's claims that part of the value of friends, with whom one shares thought and conversation, consists in their providing a "mirror" on the self (*Phaedrus* 255d5, *Nicomachean Ethics* 1169b34–35). Insofar as my friend is like me, I can appreciate my own qualities from a different perspective, which promotes my self-understanding. But there are limits to the value of mirrors. Interaction with another just like me does not itself contribute to self-criticism. This is why there is deliberative value in interaction with diverse sorts of people many of whom are not mirror images of myself. This suggests another way in which I am not deliberatively self-sufficient. Sharing thought and discussion with others, especially about how to live, improves my own practical deliberations. It enlarges my menu of options, by identifying new options, and helps me better assess the merits of these options, by forcing on my attention new considerations and arguments about the comparative merits of the options. Here we should notice the deliberative value of open and vigorous discussion with diverse interlocutors. Moreover, cooperative interaction with others allows me to participate in larger, more complex projects and so extend the scope of my deliberative control over my environment. In this way, I spread my interests more widely than I could acting on my own. Here too diversity can be helpful; cooperation is improved and extends each person further when it draws on diverse talents and skills. In these ways, interpersonal associations arguably make for fuller realization of my deliberative capacities, and this explains the prudential importance of associative relations and concern.

17. The Scope of Other- Regarding Concern

An obvious concern about this perfectionist account of other-regarding concern is its *scope*. Restricted, as it seems to be, to explaining concern for intimates, (p. 412) friends, and other associates, it seems to fall well short of the wide or universal scope of concern on which modern moral conceptions insist.

The perfectionist egoist can begin to rebut the charge of parochialism by showing just how pervasive interpersonal association is. Associations, we have seen, are not restricted to regular interactions among like-minded people. They exist whenever there is psychological continuity among people. Psychological influence can be exerted between people, on each other, even when they have not had direct interactions, as when two people influence each other through their conversations with a common third party. The ripple effects on others of our conversations, plans, actions, and relationships can extend quite widely. Moreover, continuous selves need not be connected. Any elements in a series are continuous, just in case contiguous members in the series are well connected.

This implies that noncontiguous members (e.g., the end-points) of such a series are continuous even if they are not well connected or connected at all. If so, people can be psychologically continuous who are not at all connected, provided they are members of a series of persons each of whom is connected to some degree with his neighbor in the series. There is room for debate about the comparative roles of continuity and connectedness within a psychological reductionist account of personal identity. Perhaps both relations matter and extend one's interests, but I think it is clear that continuity must matter.[25] If so, then the relations that justify other-regarding concern can extend far beyond the circle of those with whom one regularly interacts.

But can the scope of perfectionist concern be genuinely universal in scope if it is the result of interpersonal interaction? For then there must be someone—the proverbial remotest Mysian (Plato, *Theaetetus* 209b8)—with whom one has no previous relation, however indirect.[26] Should it somehow come within my power to help the remotest Mysian, at little or no cost to myself, it might seem the egoist cannot explain justified concern for him. This would represent a limitation in the scope of egoist concern.

If the remotest Mysian and I stand in no relations of connection or continuity, then his good is not already part of mine. So I can have no backward-looking reason to be concerned about him. But I can have forward-looking reasons. For it is now within my power to interact with him, and all the reasons for cultivating interpersonal association apply and provide a forward-looking rationale for concern. Even when the remotest Mysian and I have no prospect of further interaction, my assistance will enable or facilitate his pursuit of his own projects, and this will make his subsequent actions and mental states dependent upon my assistance. Indeed, other things being equal, the greater the assistance I provide, the greater is my involvement in his life. To the extent that another's actions and mental states are dependent upon my assistance, I can view the assistance as making his good a part of my own. Assistance to the remotest Mysian earns me (p. 413) a share, however small, of his good. If the perfectionist egoist can justify concern for the remotest Mysian, then the scope of such concern would seem to be genuinely universal.

18. Noninstrumental Concern for Others

But even if the scope of perfectionist concern is acceptable, perhaps its nature is not. Morality seems to require not just that we perform the actions it demands of us but also that we fulfill its demands from the right sort of motives, and sometimes morality seems to require not just that we benefit another but that we do so out of a concern for the other for her own sake. This is certainly true about the concern owed to intimates. But if justified concern for another is, as the perfectionist egoist claims, a special case of self-love, then mustn't such concern be at bottom instrumental?

The perfectionist can reply that if the good of another is a constituent part of her own good, and not just an instrumental means to the promotion of her own good, then she is justified in having intrinsic, and not merely instrumental, concern for another. When I

undergo a present sacrifice for a future benefit, I do so because the interests of my future self are interests of mine. The on-balance rationality of the sacrifice depends upon its promoting my overall good. But because the good of my future self is part of this overall good, concern for my overall good requires, as a constituent part, a concern for the good of my future self. In this way, concern for my future self for its own sake seems compatible with and, indeed, essential to self-love. If psychological relations extend an agent's interests, then the good of others can be part of my overall good just as my own future good can be. Though the on-balance rationality of other-regarding action depends upon its promoting my overall good, concern for my overall good requires, as a constituent part, concern for the welfare of those to whom I am appropriately psychologically related.

(p. 414) 19. The Variable Weight of Other-Regarding Concern

Another issue concerns the *weight* of the reasons for other-regarding concern. Both connectedness and continuity are matters of degree. If we think of degrees of connectedness and continuity in terms of a set of concentric circles, with myself occupying the inner circle and the remotest Mysian occupying the outer circle, then, as we extend the scope of psychological interdependence, the strength of the relevant psychological relations appears to weaken, and the weight of one's reasons to give aid and refrain from harm presumably weakens proportionately. Despite the wide scope of justified concern, it must apparently have variable weight. Is such an interpersonal discount rate acceptable?[27]

An interpersonal discount rate of moral concern need not be a threat to our understanding of morality. For it is commonly thought that, even if morality has universal scope, the demands it imposes are a function not simply of the amount of benefit that one can confer but also of the nature of the relationship in which one stands to potential beneficiaries. Common–sense morality recognizes more stringent obligations toward those to whom one stands in special relationships—for instance, toward family and friends and toward partners in cooperative schemes—than toward others. It seems a reasonable hypothesis that the interpersonal relationships that have special moral significance are just those relationships of psychological interaction and interdependence that extend one's interests. If so, then there will be a moral discount rate that is isomorphic to the egoist interpersonal discount rate.

I have now sketched how within a form of perfectionist egoism one might derive other-regarding concern that is both universal in scope and variable in weight. Indeed, it would seem that we have succeeded in deriving the central claims of self-referential altruism from within a purely agent-relative form of egoism that does not recognize any underived demands of impartiality. Because the perfectionist conception of prudence appears to

have normative authority, this is reason to take seriously its justification of moral concern with wide scope and variable weight.

20. Consequentialism's Prospects

If we identify consequentialism with agent-neutral conceptions of impartiality, such as utilitarianism, as many do, then I think that we must be skeptical of (p. 415) consequentialism. This sort of consequentialism cannot be the whole truth about morality. For while the agent-neutral conception of impartiality is attractive and surprisingly robust, such a conception of impartiality cannot do justice to associative obligations. These obligations resist capture within the intellectual net of agent-neutral consequentialism, and they admit of a philosophical rationale at least as plausible as anything the agent-neutral consequentialist has to offer. An adequate moral theory must recognize the demands of partiality, as well as those of impartiality.

But we can and should understand consequentialism more broadly, to include any view that takes the good to be explanatorily primary and understands other notions, such as duty and obligation, to supervene on the promotion of value. In particular, we can and should recognize forms of consequentialism that are not purely agent-neutral. One common–sensical form of consequentialism is self-referential altruism, which combines agent-neutral and agent-relative claims in a way that tries to capture both impartial and partial demands. While self-referential altruism has considerable intuitive plausibility, it leaves the normative authority of its commitment to agent-neutral impartiality unexplained. Why exactly should I be concerned about the weal and woe of others, regardless of their relationship to me? A related concern is that the self-referential altruist must answer the questions "Whom should I care about?" and "How much should I care about them?" in entirely different ways. Without some explanation of the normativity of its impartiality and some explanation for why these questions should be answered so differently, self-referential altruism may appear problematic.

These forms of consequentialism contrast with egoism, which is fully agent-relative, insisting that something is one's duty just insofar as it promotes one's own well-being or happiness. The obvious concern about egoistic consequentialism is with the stability and scope of its justification of other-regarding moral concern. These doubts are well founded when applied to forms of egoism employing traditional subjective conceptions of the good and relying on strategic arguments. Such strategic forms of egoism justify other-regarding concern that is limited in scope and counterfactually fragile. Moreover, subjective conceptions of the good fail to explain why the agent has reason to promote his own good. But a perfectionist conception of egoism fares better here. If we conceive of the good as consisting in the exercise of the very deliberative capacities that make someone an agent in the first place, we can see why he has reason to take an interest in his own good. But the relations that unify different parts of the agent's own life hold, to a significant degree, between the agent and others, which gives him reason to regard their good as part of his own good, in much the same way that the good of his future self is

part of his own good. Such claims are not unfamiliar as applied to one's relationships with one's intimate associates. But the central claims apply much more widely. Indeed, because the act of benefiting another actually constitutes association, in the relevant sense, the scope of other-regarding concern (p. 416) that can be justified in this way is genuinely universal. On this view, the strength of one's reasons to be concerned about others is proportional to the strength of the associational bonds. It follows that this kind of perfectionist egoism can justify other-regarding concern with universal scope but variable weight. This means that we can derive the attractive mix of impartiality and partiality characteristic of self-referential altruism from purely agent-relative, indeed, egoistic foundations. In doing so, we appeal to foundations whose normative authority is clearer than it was in the case of self-referential altruism itself. We also provide a unified explanation of whom to care about and how much to care. We should care about others insofar as they are or will be psychologically connected to us in the right way and to an extent proportional to the degree to which we are or will be so connected.

There remain important questions about the structure and implications of these agent-relative conceptions of consequentialism, answers to which will affect the adequacy of such conceptions. Though these issues are beyond the scope of this study, they deserve mention. The main issue concerns how to combine the interests of different people within such views. This combinatorial issue will affect how demanding such views are.

Consider self-referential altruism. It recognizes the impartial demand to be concerned about anyone whom it is in one's power to benefit but claims that the strength or urgency of one's obligation to another is a function of the nature of the relationship between benefactor and beneficiary. This would treat associative relations as putting a sort of thumb in the scales of a utilitarian calculation so as to create a normative bias for associates. On this view, an agent is required to perform that action whose value is greatest after the consequences for everyone have been recorded and multiplied by the relevant factor (equal to or greater than one) corresponding to the strength of the relationships between the agent and beneficiaries. However, until we know how great the *associate-bias* is, it is hard to know or assess the consequences of accepting self-referential altruism.

One reason utilitarianism appears to be quite demanding of some people is that the world contains a great deal of suffering, some of which can be very efficiently relieved if the better-off make sacrifices. If others are not making their share of sacrifice (partial compliance), utilitarian demands for sacrifice will apparently increase. If each of us ought to give until the point that our sacrifices are as great as the benefits we confer, then, given the conditions of partial compliance, compliers ought to sacrifice a great deal. This sort of sacrifice would involve a very significant change in lifestyle for most of those living reasonably comfortable lives and would require sacrifices that would constrain the satisfaction of their associative duties.

Would the introduction of an associate-bias significantly reduce the amount of sacrifice required? Given the very high benefit-cost ratio of many relief operations—where I can

Subscriber: King's College London; date: 14 February 2019

save many lives by very small contributions—it is difficult (p. 417) to see how an associate-bias would significantly reduce utilitarianism's demands for aid under normal conditions of partial compliance, unless the bias is very large indeed. But self-referential altruism would then lose its main appeal in relation to agent-neutral conceptions of consequentialism, inasmuch as associative duties would never trump general duties of beneficence. Alternatively, if the bias is very large and is constant across different contexts, then associational demands are likely to defeat impartial demands in all contexts, including partial compliance. The resulting view would verge on a fairly complacent moral theory that involves very little transcendence. Insofar as perfectionist egoism implies self-referential altruism, the same issues arise for it. The only difference is that with perfectionist egoism the questions concern the combination of different aspects (more and less self-confined aspects) of the agent's own good, rather than the combination of the agent's own good and that of others.

This version of the combinatorial problem arises when we allow the demands of partial and full compliance to diverge. One way for the agent-neutral consequentialist to respond is to argue that the limits of beneficence under conditions of partial compliance should be set by the amount of beneficence that would be optimal under conditions of full compliance (see Hooker, 2000; Murphy, 2000). This would reduce the demands of beneficence in conditions of partial compliance, though it is unclear how far they would be reduced. The self-referential altruist or perfectionist egoist could presumably appeal to the same device to link the demands in partial compliance to those in full compliance. Because such views already constrain the duties of beneficence by duties to oneself and one's associates, linking the demands of partial compliance to those of full compliance would make the resulting demands of beneficence even more manageable than agent-neutral consequentialism would allow.

This strategy of response to the combinatorial problem holds some promise. But it appears to depart from traditional direct, maximizing consequentialism inasmuch as it employs indirection reminiscent of rule consequentialisms. Whether such a strategy can be housed within a defensible form of consequentialism is a matter for further study. What is certain is that the combinatorial problem, the demands of morality, and the relationship between the demands of partial and full compliance are issues that any moral theory must tackle.

References

Adams, Robert. 1976. "Motive Utilitarianism." *Journal of Philosophy* 73: 467–481.

Alexander, Larry. 1987. "Scheffler on the Independence of Agent-Centered Restrictions." *Journal of Philosophy* 84: 277–283.

Annas, Julia. 1993. *The Morality of Happiness.* Oxford: Clarendon Press.

Bales, Eugene. 1971. "Act Utilitarianism: Account of Right-Making Characteristics or Decision-Making Procedures?" *American Philosophical Quarterly* 8: 257–265.

Bentham, Jeremy. [1789]. 1970. *An Introduction to the Principles of Morals and Legislation.* London: Athlone Press.

(p. 421) Boyd, Richard. 1989. "How to Be a Moral Realist." In *Essays on Moral Realism*, ed. G. Sayre-McCord, 181–228. Ithaca, N.Y.: Cornell University Press.

Bradley, F. H. [1876]. 1927. *Ethical Studies.* 2nd. ed. Oxford: Clarendon Press.

Brandt, Richard. 1963. "Toward a Credible Form of Utilitarianism." In Darwall, 2002, 207–235.

Brandt, Richard. 1979. *A Theory of the Good and the Right.* Oxford: Clarendon Press.

Brink, David O. 1989. *Moral Realism and the Foundations of Ethics.* New York: Cambridge University Press.

Brink, David. 1993. "The Separateness of Persons, Distributive Norms, and Moral Theory." In *Value, Welfare, and Morality*, ed. R. Frey and C.Morris, 252–289. New York: Cambridge University Press.

Brink, David. 1997a. Critical notice of David Lyons, *Rights, Welfare, and Mill's Moral Theory. Philosophy and Phenomenological Research* 57: 713–717.

Brink, David. 1997b. "Rational Egoism and the Separateness of Persons." In *Reading Parfit*, ed. J. Dancy, 96–134. Oxford: Blackwell.

Brink, David. 1997c. "Self-Love and Altruism." *Social Philosophy* and *Policy* 14: 122–157.

Brink, David. 2001. "Impartiality and Associative Duties." *Utilitas* 13: 152–172.

Broad, C. D. [1953]. 1971. "Self and Others." In *Broad's Critical Essays in Moral Philosophy*, ed. D. Cheney, 262–282. London: Allen and Unwin.

Darwall, Stephen. 1997. "Self-Interest and Self-Concern." *Social Philosophy and Policy* 14: 158–178.

Darwall, Stephen, ed. 2002. *Consequentialism.* Oxford: Blackwell.

Gauthier, David. 1986. *Morals by Agreement.* New York: Oxford University Press.

Gibbard, Allan. 1990. *Wise Choices, Apt Feelings: A Theory of Normative Judgment.* Cambridge, Mass.: Harvard University Press.

Green, T. H. [1883]. 2003. *Prolegomena to Ethics.* Ed. D. Brink. Oxford: Clarendon Press.

Griffin, James. 1986. *Well-Being.* Oxford: Clarendon Press.

Harsanyi, John. 1978. "Bayesian Decision Theory and Utilitarian Ethics." In Darwall, 2002, 197–206.

Hooker, Brad. 2000. *Ideal Code, Real World: A Rule-Consequentialist Theory of Morality.* Oxford: Clarendon Press.

Hurka, Thomas. 1993. *Perfectionism.* Oxford: Clarendon Press.

Huxley, Aldous. 1946. *Brave New World.* 2nd ed. New York: Harper and Row.

Jeske, Diane. 1993. "Persons, Compensation, and Utilitarianism." *Philosophical Review* 102: 541–575.

Joseph, H.W.B. 1931. *Some Problems in Ethics.* Oxford: Clarendon Press.

Kagan, Shelly. 1989. *The Limits of Morality.* Oxford: Clarendon Press.

Kamm, Frances. 1996. *Morality, Mortality.* Oxford: Clarendon Press.

Kitcher, Philip. 1999. "Essence and Perfection." *Ethics* 110: 59–83.

Kraut, Richard. 1979. "Two Conceptions of Happiness." *Philosophical Review* 88: 176–196.

Locke, John. [1690]. 1979. *An Essay Concerning Human Understanding.* Ed. P. H. Nidditch. Oxford: Clarendon Press.

Lyons, David. 1965. *Forms and Limits of Utilitarianism.* Oxford: Clarendon Press.

Lyons, David. 1994. *Rights, Welfare, and Mill's Moral Theory.* New York: Oxford University Press.

MacIntyre, Alasdair. 1981. *After Virtue.* London: Duckworth.

Mill, J. S. [1843]. 1844. *A System of Logic.* London: Longmans.

(p. 422) Mill, J. S. [1859]. 1978a. *On Liberty.* Indianapolis: Hackett.

Mill, J. S. [1861]. 1978b. *Utilitarianism.* Indianapolis: Hackett.

Moore, G. E. 1903. *Principia Ethica.* Cambridge: Cambridge University Press.

Mueller, Dennis. 1979. *Public Choice.* New York: Cambridge University Press.

Murphy, Liam. 2000. *Moral Demands in Nonideal Theory.* Oxford: Clarendon Press.

Nagel, Thomas. 1970. *The Possibility of Altruism.* Oxford: Clarendon Press.

Nagel, Thomas. 1986. *The View from Nowhere.* New York: Oxford University Press.

Nozick, Robert. 1974. *Anarchy, State, and Utopia.* New York: Basic Books.

Nozick, Robert. 1980. *Philosophical Explanations.* Cambridge, Mass.: Harvard University Press.

Parfit, Derek. 1984. *Reasons and Persons.* Oxford: Clarendon Press.

Pettit, Philip. 1991. "Consequentialism." In Darwall, 2002, 95–107.

Rae, Douglas. 1969. "Decision-Rules and Individual Values in Constitutional Choice." *American Political Science Review* 63: 40–56.

Railton, Peter. 1984. "Alienation, Consequentialism, and the Demands of Morality." *Philosophy* and *Public Affairs* 13: 134–171.

Railton, Peter. 1986. "Facts and Values." *Philosophical Topics* 14: 5–31.

Rawls, John. 1971. *A Theory of Justice.* Cambridge, Mass.: Harvard University Press.

Reid, Thomas. [1785]. 1969. *Essays on the Intellectual Powers of Man.* Ed. B. Brody. Cambridge: MIT Press.

Rosati, Connie. 1995. "Persons, Perspectives, and Full Information Accounts of the Good." *Ethics* 105: 296–325.

Ross, W. D. 1930. *The Right and the Good.* Oxford: Clarendon Press.

Sandel, Michael. 1982. *Liberalism and the Limits of Justice.* Cambridge, Mass.: Harvard University Press.

Scanlon, T. M. 1982. "Contractualism and Utilitarianism." In *Utilitarianism and Beyond*, ed. A. Sen and B. Williams, 103–128. New York: Cambridge University Press,

Scanlon, T. M. 1998. *What We Owe to Each Other.* Cambridge, Mass.: Harvard University Press.

Scheffler, Samuel. 1982. *The Rejection of Consequentialism.* Oxford: Clarendon Press.

Scheffler, Samuel. 1992. *Human Morality.* New York: Oxford University Press.

Scheffler, Samuel. 1995. "Families, Nations, and Strangers." The Lindley Lecture. Lawrence: Department of Philosophy, University of Kansas.

Scheffler, Samuel. 1997. "Relationships and Responsibilities." *Philosophy* and *Public Affairs* 26: 189–209.

Shoemaker, Sydney. 1963. *Self-Knowledge and Self-Identity.* Ithaca, N.Y.: Cornell University Press.

Shoemaker, Sydney. 1984. "Personal Identity: A Materialist's Account." In S. Shoemaker and R. Swinburne, *Personal Identity*, 67–132. Oxford: Blackwell.

Sidgwick, Henry. [1907]. 1966. *The Methods of Ethics.* New York: Dover.

Simmons, A. John. 1996. "Associative Political Obligations." *Ethics* 106: 247–273.

Slote, Michael. 1985. *Common-Sense Morality and Consequentialism.* London: Routledge and Kegan Paul.

Sobel, David. 1994. "Full Information Accounts of Well-Being." *Ethics* 104: 784–810.

Taylor, Charles. 1989. *Sources of the Self.* Cambridge, Mass.: Harvard University Press.

Taylor, Michael. 1969. "Proof of a Theorem on Majority Rule." *Behavioral Science* 14: 228–231.

Walzer, Michael. 1983. *Spheres of Justice.* New York: Basic Books.

(p. 423) Wiggins, David. 1967. *Identity and Spatio-Temporal Continuity.* Oxford: Blackwell.

Williams, Bernard. 1973. "A Critique of Utilitarianism." In J.J.C. Smart and B. Williams, *Utilitarianism: For and Against*, 75–155. New York: Cambridge University Press.

Williams, Bernard. 1976. "Persons, Character, and Morality." In Williams, 1981, 1–19.

Williams, Bernard. 1981. *Moral Luck.* Cambridge: Cambridge University Press.

Notes:

(1.) If good character is itself explained in terms of promotion of value, then this sort of virtue theory would itself be a special kind of consequentialist view, namely, motive consequentialism.

(2.) This defense of maximization is act consequentialist. Mill's view here is interesting. While many claims in *Utilitarianism* point toward act utilitarianism, an important strand in Mill's utilitarianism actually defines right action in terms of the utility of blaming the conduct in question. In particular, he claims that one is under an obligation to do something just in case failure to do it is wrong and that an action is wrong just in case some kind of external or internal sanction—punishment, social censure, or self-reproach—ought to be applied to its performance (1978b, ch. 5, para. 14). Whether sanctions ought to be applied to an action (hence whether it is wrong) depends on the utility of doing so (1978b, ch. 5, par. 25). This strand in Mill's theory ties wrongness to blame in a way that act utilitarianism does not. These aspects of Mill's theory are discussed in Lyons, 1994, and Brink, 1997a.

(3.) The issues here are complicated. In some such cases, it does seem pointless or even pernicious to conform to rules that would have optimal general acceptance value when others are not so conforming. But the appeal to general acceptance value may be an advantage in other cases of partial compliance. See the discussion in section 20.

(4.) Williams's concern is with impartial moralities, including, but not restricted to, utilitarianism. However, the worry seems to apply to a variety of optimizing theories, not just impartial conceptions.

(5.) Parts of this account of having one's conduct regulated by a principle, without constantly consulting it, can be found in Mill's claims in *Utilitarianism* (1978b, ch. 2, pars. 19, 23–25) and *A System of Logic* (1844, VI.xii.7) about the need for "secondary principles" that function in our practical reasoning in lieu of direct appeals to the utilitarian first principle. Scheffler, 1992, ch. 3, contains a nice discussion of the variety of roles that moral principles can play in moral deliberation.

(6.) So I do not distinguish, as some do, between happiness and well-being. Some distinguish the two, because, whereas they can entertain objective conceptions of well-being, they regard happiness as an inherently subjective concept that does not admit of objective conceptions. I am not persuaded of this contrast; I think that we can entertain and take seriously objective (including perfectionist) conceptions of happiness (see Kraut, 1979). Anyone who disagrees and thinks that subjectivism about happiness is true by definition can simply put happiness to the side and reinterpret the discussion solely in terms of well-being.

(7.) Thomas Hurka, chapter 13 in this volume, provides a more systematic discussion of the good and defends some related conclusions.

(8.) However, it should perhaps be noted that hedonism is objective, insofar as it claims that pleasure is a person's good whether or not the person realizes this or desires pleasure.

(9.) Ideal preference views sometimes trace their ancestry to Mill's appeal in *Utilitarianism* to the preferences of competent judges to identify higher pleasures, though Ithink that Mill's claims can be given a consistent perfectionist reading. Important statements of the informed preference view include Brandt, 1979, and Griffin, 1986. The most sophisticated version of the view of which I am aware is Railton, 1986.

(10.) Moore endorses an objective list, 1903, ch. 6, as does Ross, 1930, p. 140. Parfit discusses such theories sympathetically (1984, pp. 493–502).

(11.) This is like the criticism, made by Joseph, among others, that the intuitionist's objective list of right-making factors amounts to nothing more than an "unconnected heap" of obligations (see Joseph, 1931, p. 67).

(12.) A vigorous contemporary statement of perfectionism is Hurka, 1993.

(13.) Kitcher, 1999, raises some relevant difficulties for Hurka's appeal to a biological essence.

(14.) Whereas unanimity may be the only decision rule acceptable to all ex post, majority-rule can be acceptable to all ex ante (see Mueller, 1979, ch. 11; Rae, 1969; M. Taylor, 1969). Harsanyi offers such a contractualist defense of utilitarianism (1978). For more discussion of the compatibility of contractualism and utilitarianism, see Brink, 1993.

(15.) Even Kamm does not think inviolability should be absolute (1996, 2: 274).

(16.) These thoughts may also suggest a reply to Rawls's argument that utilitarianism violates the strains of commitment within a well-ordered society (1971, esp. pp. 175–183, 496–502). If a well-ordered society is one in which citizens are regulated by a sense of justice, informed by a utilitarian conception of impartiality, then utilitarianism may not impose undue strains of commitment.

(17.) For communitarian conceptions that tie the content, as well as the ground, of associative duties to the terms of past association, see Bradley, 1927, esp. ch. 5; MacIntyre, 1981, esp. ch. 15; Sandel, 1982; C. Taylor, 1989; Walzer, 1983. The fact that my conception does not tie the content of associative duties to past association allows me to avoid Simmons's worries about grounding special obligations in morally imperfect associations (see Simmons, 1996, esp. p. 266).

(18.) This way of dealing with morally imperfect associations does not restrict the kind of interdependence and influence that generates associative duties but insists that associative duties enjoin concern for the associate's well-being. Alternatively, one might try to restrict the kind of interdependence and influence that generates associative duties in the first place, so that certain kinds of morally objectionable forms of association do not generate associative duties at all. One might insist that some degree of cooperative interaction and good will are essential ingredients of normatively significant association, much as some degree of diachronic cooperation and good will are arguably essential to unity within a single life. But just as intrapersonal unity must be compatible with some changes of mind and heart, so, too, normatively significant forms of interpersonal association cannot be limited to the virtuous. But then it becomes difficult to know how to restrict the normatively significant forms of association. However, this alternative deserves further study.

(19.) However, the intrinsic normative significance of special relations cannot be captured by recognizing the intrinsic value of associative relations within a utilitarian view. For instance, the utilitarian can assign special intrinsic value to friendship. But this won't allow her to claim that an agent has reasons to give priority to his own friend when he could provide comparable or greater benefits to the friend of someone who is a perfect stranger to him.

(20.) See Scheffler, 1995, 1997. In the latter work, Scheffler offers his own rationale for associative duties, which I won't discuss here.

(21.) Interestingly, whereas Kagan recognizes that any moral conception, including utilitarianism, must meet the demand for a philosophical rationale, he presses this demand only against friends of constraints and options, not against utilitarianism itself. See Kagan, 1989, pp. 18–19.

(22.) If we are to define identity, which is a one-to-one relation, in terms of psychological continuity, which can take a one-to-many form, we must define it in terms of *nonbranching* psychological continuity. But the reasoning that leads us to this conclusion

Subscriber: King's College London; date: 14 February 2019

may also lead us to the conclusion that it is continuity, rather than identity per se, that is what has primary normative significance. I discuss these matters in Brink, 1997b, 1997c.

(23.) Branching cases, such as fission, in which consciousness divides and psychological continuity is maximal but takes a one–many form, represent the limiting case of interpersonal psychological continuity. In fission cases, there is, by hypothesis, no less continuity than in normal intrapersonal cases. What makes the former interpersonal is simply that in them continuity takes a one–many form. Insofar as our primary concern is with psychological continuity, whether or not it takes a unique or nonbranching form, fission cases throw further doubt on the assumption that the separateness of persons is fundamental. See Brink, 1997c, pp. 138–143.

(24.) This is an act egoist justification of other-regarding moral norms. But its nature and limits bear comparison with Gauthier's (1986) motive egoist justification.

(25.) Continuity must figure in a reductionist account of identity if only to meet Reid's demand that any criterion of identity be transitive (see Reid, 1969, p. 358).

(26.) The introduction of the proverbial remotest Mysian into discussions of the scope of ethical concern is discussed by Annas, 1993, ch. 12.

(27.) The precise shape of the interpersonal discount rate is a matter for further investigation. As long as psychological continuity is one of the relations that matter, a significant threshold of concern can be justified well out into outer circles. But as long as psychological connectedness is also one of the psychological relations that matter, there will nonetheless be significant differences in the degree of concern that can be justified, above this threshold, in different circles, because an agent will be differentially psychologically connected to others.

David O. Brink

David O. Brink is Professor of Philosophy at the University of California, San Diego and Director of the Institute for Law and Philosophy at the University of San Diego School of Law. His research interests are in ethical theory, history of ethics, and jurisprudence. He is the author of Moral Realism and the Foundations of Ethics (Cambridge, 1989) and Perfectionism and the Common Good: Themes in the Philosophy of T.H. Green (Oxford, 2003).

Subscriber: King's College London; date: 14 February 2019