You only need to read the sections with the highlighted heading (1, 2, 3, 3.2, 3.3, 4, 4.1, 5)

Abstract and Keywords

'Normative ethics' is an enormous field. It is concerned with the articulation and the justification of the fundamental principles that govern the issues of how we should live and what we morally ought to do. Its most general concerns are providing an account of moral evaluation and, possibly, articulating a decision procedure to guide moral action. Though both these aims rely on articulating the correct set of moral principles that govern evaluation and that can also be used in articulating a decision procedure or rule, they are not coextensive. Recent critical work, especially on the part of particularists and virtue ethicists, has generated more pressure to separate clearly the two.

Keywords: normative ethics, moral evaluation, moral action, decision procedure, moral principles, virtue ethics

1. Introduction

'Normative ethics' is an enormous field. It is concerned with the articulation and the justification of the fundamental principles that govern the issues of how we should live and what we morally ought to do. Its most general concerns are providing an account of moral evaluation and, possibly, articulating a decision procedure to guide moral action. Though both these aims rely on articulating the correct set of moral principles that govern evaluation and that can also be used in articulating a decision procedure or rule, they are not coextensive. Recent critical work, especially on the part of particularists and virtue ethicists, has generated more pressure to clearly separate the two.[1]

More specific concerns have to do with investigating the nature of value, for example, and related issues of weighing goods, incommensurability, and the nature of virtue and vice. Indeed, some of these projects, which began as ones peripheral to an understanding of moral evaluation and action guidance, have taken on a life of their own and have been responsible for generating connections between normative ethics and epistemology, economics, and psychology.

Normative ethical theories have traditionally been divided into teleological or deontological categories. Teleological theories are thought to be those that define moral quality in terms of the achievement of some good or avoidance of some bad. Thus, for these accounts, a theory of value is extremely important because it will (p. 32) provide substance to what the theory recommends. Deontological accounts are those that specify moral quality—not in terms of goodness achieved, but rather as a function of something else, such as a duty that binds, no matter the consequences, or the adherence to principles that are agreed upon by rational agents. This is not to say that an account of value is not important to deontological approaches—it is. But these approaches do not define the right in terms of the good, so to speak.[2].

2. Value

Any teleological account of right or virtuous action will need to develop a theory of value —that is, a theory of the good that is to be caused or accomplished by the action or person. It is important to keep in mind that the theory of value, however, is only a portion of a full-fledged normative ethical theory. Such theories of value, for example, can also be appealed to in accounts of prudence or rationality.

In normative ethics, well-being is often taken to be the ultimate good for a person. It had been hoped that an elegant, simple, unified theory of value could be developed. However, it is possible that any kind of truly unified value theory is hopeless. As Thomas Nagel writes: 'I do not believe that the source of value is unitary—displaying apparent multiplicity only in its application to the world. I believe that value has fundamentally different kinds of sources, and that they are reflected in the classification of values into types. Not all values represent the pursuit of some single good in a variety of settings' (1979: 131–2). As we shall see, there are good reasons for Nagel's pessimism. A fragmented account of value would have serious implications for any ethical theory, and also have implications for views on moral dilemmas.[3] Nagel believes that not only are values plural, they are incommensurable or 'incomparable'. If so, that would mean that some moral dilemmas have no resolution.

Yet, a variety of unified theories of value have been proposed, the most well known being hedonism and desire satisfaction accounts of "good" and "good for".

Though few currently accept evaluative hedonism as the correct theory of value, it is certainly worth discussing in this context since most writers define their views relative to a rejection of hedonism. Hedonism is the view that pleasure is the only intrinsic good and pain the only intrinsic bad.[4] A number of writers have presented (p. 33) compelling counter-examples to this claim. Nozick's 'experience machine' example asks us to consider the following choice (Nozick 1974: 42–5): Suppose one could plug into a machine that could generate pleasurable experiences for the rest of one's life. Thus, one could believe one was actually doing whatever one wanted—sitting on a beach in Cancun, eating a great meal in Paris, or winning the Nobel Peace Prize. Would it be "good for" one to choose to be plugged into this experience machine, even assuming that it will guarantee a

greater balance of pleasurable experiences through one's life? Nozick, of course, believed not. And many others have the same intuition: experiences that are veridical (reflect reality) have greater value than experiences that are merely pleasurable. Thus, an objective feature seems relevant to the value of the experience.

Theoretical advantages to hedonism were thought to involve its relative simplicity: with just one basic good, the problem of weighing incommensurable goods can be avoided. However, even hedonism, which is relatively simple, cannot avoid the problem altogether. Most accounts of pleasure note that pleasure can be measured along certain distinct parameters—intensity and duration are just two examples.[5] The weighing problem will emerge in comparing pleasures along varying parameters of measurement.[6] The problem of interpersonal comparisons of "utility" occurs here as well.[7]

Thus, the theoretical advantage is not as profound. Given the counter-intuitive aspect of this account of value, alternatives that are more complicated do a better job of conforming to our considered judgements. There is by no means universal agreement on this, however. For example, Fred Feldman attempts to develop a more sophisticated version of hedonism, propositional hedonism, which is an improvement over the classical view.[8] Also, some writers have noted that Nozick's experience machine thought experiment gets its force from intuitions formed under normal circumstances, and those circumstances do not include experience machines (Railton 1989).

Other subjective or quasi-subjective accounts replace pleasure with preference or desire satisfaction. But these approaches also suffer from the problem that what we (p. 34) seem to find valuable is objective states of affairs existing independently. Thus, one might believe that certain aesthetic features, like beauty, have value even in the absence of any appreciation of beauty (Moore 1988). There is the additional problem that desire satisfaction views tend to reverse the connection between what is good and what is desired. The intuition is that x is not good for a because a desires x; rather, a desires x because x is good for a.[9]

These considerations have led some to reject the view that well-being is captured by appeal to subjective states—and thus reject both hedonism and desire or preference satisfaction approaches to value—in favour of some sort of objective list theory. Some things are good even if they don't cause pleasure, and are independent of our desires and/or preferences. Candidates will include not only beauty, but also things like knowledge. How do we decide what goes on this list? Recently, Tom Hurka (1993) has suggested a form of perfectionism. It is those things that constitute perfections of human nature that are constituents of human well-being. One might argue, then, that it is in our nature to seek knowledge, for example, and develop our rational capacities. Thus, knowledge forms a part of our well-being. Note that on this view, however, it is possible for something to be good for one even if one does not want it or even if it would not lead to one experiencing a pleasurable response.

Since some argue that the very point of morality is well-being, this issue is of enduring interest. It is of crucial importance, of course, to the so-called "teleological" accounts of moral evaluation, which we turn to next.[10]

3. Consequentialism

Consequentialist moral theories are those that determine the rightness of an action, appropriateness of blame, and so on, solely by a consideration of the consequences generated by the action in question.[11] The most prominent version of consequentialism is utilitarianism, which holds that the right action maximizes the good.

(p. 35)

Classical act utilitarianism held that the right action maximized pleasure. While utilitarianism dominated analytic normative ethics for most of the last century, few modern writers accepted the classical version of the theory as proposed by Bentham and Mill in the nineteenth century. Problems with their commitment to hedonism led to a rejection of the value theory (discussed in the previous section). However, aside from this concern, problems were also raised for the theory's commitment to maximization or *promotion* of the good.[12] For example, utilitarianism seems insensitive to distribution issues and the distinctness of persons; it fails to reflect our intuitions regarding desert; it is too demanding; and it has been accused of ignoring agent-relative reasons through a commitment to agent-neutrality.[13] Each of these criticisms has prompted a response that has led to a modification.

One problem is that of understanding what, exactly, the principle of utility will recommend in terms of aggregating the good. A question arises as to whether the theory recommends that we strive to maximize average utility or total utility. The contrast is noted by a number of writers. Smart summarizes the choice this way:

> Would you be quite indifferent between (a) a universe containing only one million happy sentient beings, all equally happy, and (b) a universe containing two million happy beings, each neither more nor less happy than any in the first universe? Or would you, as a humane and sympathetic person, give a preference to the second universe? (Smart and Williams 1973)

Smart himself favours the second, where the total is greater, though he thinks that in point of fact the distinction will have little practical import. Other writers, however, are not as optimistic. Total utilitarianism runs the risk of endorsing Parfit's (1984) famous "repugnant conclusion"—that is, it would seem committed to recom-mending a world with millions of marginally happy people over one with a much smaller group of quite happy people, because the total amount of happiness would be greater. But the average option, which avoids this implication, at least superficially, may run afoul of "mere addition".[14]

Another problem is the issue of justice or fairness. This problem comes up in a variety of contexts, and there are numerous examples offered to show that (p. 36) utilitarianism is flawed because it will recommend to the agent that she perform acts that are obviously

unjust. The classic example is this: Suppose that, to stop a spate of horrific murders, it proves necessary to frame an innocent person for one of the crimes in such a way that all would be convinced. If this act deterred future crimes, then it looks as though the utilitarian must recommend that the innocent person be sacrificed to save the other innocent persons.[15] Yet this seems quite incompatible with justice. Of course, to the utilitarian, if the act is what promotes the good, then it is right, so it would be a misdescription to say that the utilitarian tells us to do what is wrong if that will maximize the good.[16]

Utilitarianism is also criticized for its demandingness. For example, it would seem that an agent's time would be better spent working for Oxfam as opposed to taking a vacation. Thus, taking the vacation is wrong. This issue has generated a good deal of anti-consequentialist sentiment. The problem, again, is with the *maximizing* or *optimizing* element of the theory. Given that *x* is good, it does, intuitively, seem plausible to hold that more of *x* is better than less. Indeed, that seems only rational.

Some utilitarians simply bite the bullet, Singer (1972) and Kagan (1989) being examples of those who hold that morality just *is* demanding. Wishing it weren't so doesn't make it so. Those searching for a justification for a suburban-style ethics are searching in vain.[17] It should be noted as well that this concern is related to a separate criticism, that utilitarianism makes morality *allpervasive*.For example, how can eating an apple as opposed to an orange be a *moral* issue or choice? This sort of challenge is distinct from both the overriding and the demandingness issues.

(p. 37)

Many find the demandingness of utilitarianism problematic and try to offer alternative consequentialist views. For example, satisficers can avoid the problem by jettisoning the maximizing condition.[18] And, because satisficers can also be consequentialist, it seems mistaken to equate consequentialism with maximizing views. But the satisficing approach suffers from severe problems; namely, as noted above, it just seems irrational to prefer less good to more. The sorts of cases used to motivate the view can be readily explained as ones that are oversimplified. If ice cream is good, but more ice cream can sometimes be worse than less, then that shows that maximizing can't be right. But really all that shows is that maximizing ice cream can't be right. Ice cream is not a final good. And, as Philip Pettit points out, a consequentialist can motivate some 'satisficing' policies—ones that promote efficiency, for example. But 'unmotivated sub-optimizing' does seem quite irrational.[19]

One troubling aspect of demandingness is that the demand made on the individual in light of non-compliance by others—indeed, non-compliance by the vast majority of others—can be very weighty. It seems unjust that I should have to donate all of my assets to famine relief to help save starving people when it is also true that if everyone donated a bit of their income the starvation problem would be greatly alleviated. Why should my responsibility be greater simply because others fail to live up to theirs? Liam Murphy has suggested a 'collective principle of beneficence' that would avoid the problem by recommending that everyone do what is optimal of the options available, except in situations of non-compli-

ance. That is, in situations of partial compliance 'a person's maximum level of required sacrifice is that which will reduce her level of expected well-being to the level it would be, all other aspects of her situation remaining the same, if there were to be full compliance from that point on'. This view, which presents a limited constraint on maximizing, clashes with our intuitions about rescue situations, as Murphy himself notes. If I can rescue more than one person, but my "fair share" given that others are present and also able to help is only one, it looks like Murphy says that I've fulfilled my obligation after the first rescue. But it seems counter-intuitive to say that I have lived up to my obligations, though I let people drown whom I could have easily saved (Murphy 2000: 117).[20]

This concern can be expanded upon by noting that utilitarians are committed to negative responsibility. That is, intuitively, one can be held responsible for what happens as a result of one's failure to act. Bernard Williams is the writer most associated with this criticism of utilitarianism. Negative responsibility becomes problematic, on his view, when another agent is in a position to affect the morality (p. 38) of one's actions. He asks us to imagine a scenario in which Jim, travelling through a remote village, comes across an evil dictator whose captain, Pedro, plans to execute twenty innocent villagers. Jim is offered the opportunity to kill one villager, in which case Pedro will spare the remaining nineteen. If, however, Jim refuses, all twenty will be executed. While it *may* be true that Jim is obligated to kill, Williams baulks at the view that if Jim fails to kill the one, he is thereby responsible for the death of the nineteen. True, Jim could have prevented their deaths, but the responsibility rests with Pedro. As problematic as this may seem, it is important to note that our intuitions in these sorts of cases seem to vary widely. Even something like changing the perspective from which the story is told has this effect. For example, one of the villagers whose life is at stake might well feel that Jim is responsible, at least in part. Indeed, we have these intuitions in structurally similar cases in which the agent is in a position to save a life by doing something that would normally be considered immoral, such as telling a lie. If an agent could save a life by telling a lie and does not do it, then it would seem appropriate to hold him responsible, at least in part.

Within consequentialism perhaps the most promising way to avoid the demandingness problem and related problems is to develop either an indirect form of the theory, such as rule consequentialism or motive consequentialism, or an objective view, such as that suggested by Railton.[21]

3.1 The Two-Levels View

R. M. Hare suggested one consequentialist approach that was intended to accommodate our intuitions about fairness as well as special obligations. Hare distinguishes two levels of moral thinking: the level of theory (or the critical level) and the level of actual practice (or the intuitive level). The intuitive level is the level of rules the agent follows, by and large; the critical level is the explicitly act utilitarian level. Which rules produce the best acts? They are those rules that are not too complicated or too hard to inculcate and teach. They guide us in our day-to-day practical decision-making, but they are by no means absolute. Instead, they function like overridable heuristics. There will be situations

in which the agent must appeal to the critical level. For example, when selecting the principles to follow at the intuitive level—that is, the ones to inculcate and teach—we must try (p. 39) to operate at the critical level. Also, rules at the intuitive level may conflict, or an odd case may not be clearly covered by a rule, or there may be a situation in which there would be some massive loss of utility if the rule were followed, and so on. So, we must sometimes step back and think critically. Thinking at this level is to think as the ideal utilitarian would—the level of what Hare calls 'the Archangel'. The Archangel is not biased or prejudiced, and has full information and perfect reason. The truly right act is the one the Archangel would advise. However, in real life there are no archangels, so ordinary people must make do with the intuitive level, though, again, there will be circumstances in which it pays to try to adopt the critical level.

Hare argues, in a manoeuvre now fairly familiar to consequentialists, that efficiently acting at the intuitive level or efficient practical deliberation will involve the agent having the appropriate character and dispositions. These will be the dispositions, or virtues, that will enhance the agent's chances of acting optimifically. He makes this manoeuvre as a way of trying to avoid the self-defeating objection and the schizophrenia objection (described later).[22]

Hare holds that the cost of internalizing principles must be factored into overall efficiency considerations, and there are costs associated with violating rules that are effective when inculcated (at the intuitive level). Indirect versions of this theory such as rule utilitarianism also consider this factor.

3.2 Rule Consequentialism

One initially appealing strategy for dealing with the problems of fairness, demandingness, and conflicts with intuitions regarding justice is to argue that right action should be understood in terms of *rules* that maximize the good. Thus, the right action would be that action (among the alternatives) performed in accordance with the set of rules that maximizes utility. One could then argue that the utilitarian is *not* committed to the rightness of unjust acts. The reason why handing over an innocent person to a mob is wrong is because such an act conflicts with a rule that is part of a set of rules that maximize utility. Brandt (1979) nicely developed this alternative. However, the standard objection to this approach is that it is either guilty of absolutism, and thus shares familiar problems with Kant's ethics, or collapses into act utilitarianism.[23] More recently, Brad Hooker has developed an (p. 40) alternative that he hopes will avoid these problems. On his view, a *wrong* action is defined in the following way:

> An act is wrong if and only if it is forbidden by the code of rules whose internalization by the overwhelming majority of everyone everywhere in each new generation has maximum expected value in terms of well-being (with some priority for the worst off). The calculation of a code's expected value includes all costs of getting the code internalized. If in terms of expected value two or more codes are

better than the rest but equal to one another, the one closest to conventional morality determines what acts are wrong. (Hooker 2000)

Hooker believes that his approach avoids the above-mentioned dilemma for rule consequentialism by building in internalization of the rules. The optimal system would not simply reduce to 'Maximize the good' for two reasons. First, it is too vague. Secondly, if persons came to believe that this was the one and only operative rule, they would lose confidence in the capacity of others not to steal, to keep their promises, and so forth. This system avoids rigidity with a 'prevent disaster' rule that will compete with the other rules and allow for some rules to be abandoned when there is a conflict with the directive to 'prevent disaster'.

Of course, the act consequentialist need not deny rules a very significant place in morality. But those rules are rules of thumb that may be overridden by a consideration of consequences. The act utilitarian could argue that Hooker's escape clause of 'prevent disaster' simply performs basically the same overriding function. While this may make his account superior to a standard rule utilitarian view, it doesn't mark an improvement over act consequentialism.

3.3 Objective Consequentialism

In 'Alienation, Consequentialism, and the Demands of Morality' (Railton 1984) Peter Railton introduced a distinction between subjective and objective consequentialism:

> *Subjective consequentialism* is the view that whenever one faces a choice of actions, one should attempt to determine which act of those available would most promote the good, and should then try to act accordingly…*Objective consequentialism* is the view that the criterion of the rightness of an act or course of action is whether it in fact would most promote the good of those acts available to the agent.

The subjective consequentialist holds that the moral agent must consciously try to maximize the good, whereas the objective consequentialist need not recommend a particular mode of decision-making since all that matters is actually getting it right, whatever decision procedure is used. Thus, if one opts for the objective formulation of the theory, one can solve a host of problems for the view. The theory is not weirdly self-defeating.[24] It allows the theory to bypass (p. 41) Michael Stocker's criticism, levelled against ethical theories more generally, that they are schizophrenic since the agent needn't be directly motivated by the value that itself justifies the motivation. Further, a variety of permissions and constraints can be justified on consequentialist grounds. That is, as Railton notes, it may well be permissible for a husband to spend money to see his wife who is away from home rather than send the money to Oxfam, because partial concern for one's spouse is something that promotes overall well-being. We want people, in general, to have such a partial concern because it is something that, given the facts of human psychology, makes for happier people. One gives an impartial justification for partial concern.[25] This also offers at least a partial solution to the demandingness problem. I needn't give all my assets to Oxfam if, incompatible with partial concern, that would have a negative impact on those 'near and dear'.

4. Deontological Ethics

Deontological approaches are standardly contrasted with consequentialist approaches in that the right is not defined in terms of the good. Instead, while consequences may or may not be *relevant* to a determination of rightness, consequences are not the sole determinant. Further, the deontologist will recognize certain "agent-relative" factors or elements as relevant to determinations of rightness. These features are often said to contrast with consequentialism, though I view this issue as irrelevant in marking the distinction since constraints and options can be given a consequentialist justification.[26] Thus, the range of theories covered by this category is quite broad: Kantian ethics, intuitionistic approaches, and contractarianism and contractualism. The natural law approach offers another alternative, and some writers—such as Bernard Gert, who develops a Hobbesian approach to morality—view themselves within this tradition.[27]

(p. 42)

4.1 Kantian Ethics

While there is some debate about whether Kant truly counts as a deontologist, his work is generally classified as deontological since he holds that a person's duty or moral obligation is determined by adherence to a given principle that does not appeal for its justification to actual or expected consequences to be produced.[28] Kant's view is notoriously contrasted with consequentialism. Kant believed that the morally appropriate course of action must pass the test of the Categorical Imperative. There is in Kant's ethics the idea that the fundamental basis for moral duty is that there is some end that has *unconditional* value, and, for him, what has unconditional value is rational nature. Rational nature is an "end in itself". Rational beings are autonomous beings, ones free to choose actions and thus free to follow rules: their behaviour, unlike that of non-rational animals, is free and not determined by instinct or mere desire. Rational beings can follow the Categorical Imperative. Only rational beings can be moral agents, and, more controversially, only rational beings are the sorts of beings that can be *acted on* morally or immorally. A non-rational being, on Kant's view, deserves no moral respect.

The Categorical Imperative itself has several formulations, though the two most usually discussed are (1) the universalizability formulation, 'Act only according to that maxim by which you can at the same time will that it should become a universal law' (Kant 1969: 44), and (2) the principle of humanity, 'Act so that you treat humanity, whether in your own person or in that of another, always as an end and never as a means only' (Kant 1969: 54).

The Categorical Imperative is contrasted with hypothetical imperatives that specify what one ought to do given that one has certain goals. The binding force of the Categorical Imperative does not depend on the agent's goals or desires. Space precludes a detailed discussion of the fairly standard problems raised in connection with understanding the content of the formulations and applying them. However, significant questions arise in con-

nection with formulation of the maxim tested against the Categorical Imperative. Is this an all-pervasive procedure— need I be constantly testing maxims to be sure my behaviour conforms to the Categorical Imperative? If so, Kantian ethics would in its own way be very demanding.

Barbara Herman argues that it isn't specific maxims that get tested against the Categorical Imperative; rather, it is "generic" maxims that, having met the demands (p. 43) of the Categorical Imperative, establish a presumption against a certain type of act (though this presumption can be defeated by some other consideration):

> The idea is that we are to think of the CI [Categorical Imperative] procedure as applying not to actual maxims of action but to a type of action–justification pair: to do *x*-type action for *y*-type reason. I call these pairs 'generic maxims.'…The rejection of a generic maxim by the CI procedure shows that a certain kind of action may not be done for a certain kind of reason. This in effect establishes a principle of moral judgment that can set terms for moral deliberation. (Herman 1993: 147)

However, Kantian ethics, like utilitarianism, has come under attack from critics for its impartiality and oversimplification. Further, Kant's ethics—in contrast to virtue ethics, for example— seems cold and antiseptic. To Kant, the emotions have no intrinsic worth or moral significance in ethics, in deference to the role of reason and duty, whatever one's feelings. This line of attack has been pushed by writers such as Michael Stocker. Stocker (1976, 1990) presents the case of a person who is called upon to visit a sick friend in the hospital. On the Kantian view, it looks as though the morally appropriate motive for visiting one's friend would be duty, yet this seems to conflict with the ideal of friendship.

Marcia Baron, however, defends the Kantian approach here, noting that Kant's emphasis on the value of acting from duty should be seen as an acknowledgement or endorsement of the view that one's conduct should be regulated by a commitment to do what morality demands of one. She believes that the value of duty lies in its regulative role (Baron 1995: 118 ff.). It is misleading to interpret Kant as presenting an all-pervasive morality.

Virtue ethicists have also picked up on this feature of Kant's ethics, noting that Kant seems to put significance on strength of will as a primary virtue, whereas someone with a more Aristotelian view might argue for temperance.[29] Kant's unhappy agent, who must force himself to act well, contrasts with Aristotle's perfectly virtuous person, whose desires and moral beliefs are in harmony. Kant seems to hold that the continent person is a prime example of a morally good person. This feature of Kant's ethics has generated a large literature.[30] More recently, writers have been trying to argue that Kant and Aristotle are more similar on the topic of virtue than they might initially appear.[31] Intuitions in this regard may simply turn on how or what we are evaluating. For example, a common complaint against the Kantian view is that it seems to hold as a paradigm someone with unsavoury appetites who nevertheless manages to control them. The contrary intuition is: isn't it better, *pace* Aristotle, just not to have those appetites in the first place? But the question of who is the better person is distinct from the question of who has an important

quality that we can admire: we may admire will-power while also valuing the people who have no need of it.

(p. 44)

Even more problematic is the seemingly absolutist quality of Kant's ethics. He notoriously holds lying to be immoral even for altruistic motives and for an excellent end, such as to save another's life. Christine Korsgaard, however, has argued (1986) that Kant's view would offer us a plausible way to deal with evil.

One telling contrast between Kant and the utilitarians on value can be encapsulated by their account of the status of animals. On a Kantian view, animals have no independent moral status whatsoever. On the classical utilitarian view they do, simply in virtue of being sentient—that is, in virtue of possessing the capacity for experiencing pleasure and pain. This is not to say that Kant believes that "anything goes" with respect to animals: if one is cruel to them, then one runs the risk of corruption. One may become more likely to engage in cruel acts against other persons.[32] Thus, we have an indirect duty to be kind to animals, since this will reinforce our duties to other persons. This instrumentalist rationale for kindness to animals, which, incidentally, looks rather consequentialist in form, strikes many as unsatisfactory. They regard cruelty to an animal as something morally problematic because it harms *the animal*, not simply because it may lead to further harms for human beings. Kant's justification seems to miss the point.

However, one significant theoretical advantage of the internalist strategy pursued by Kant is its strategy for avoiding moral luck problems. Writers such as Thomas Nagel and Bernard Williams have highlighted the problem. Nagel discusses the case of a truck driver who runs over a child, and who is blamed, contrasted with a case that is identical except for the fact that the truck driver, through good luck, happens not to run over anyone, and thus is not blamed. Luck is the only difference, yet our moral practice is such that we do seem to regard the first case as worse than the second. On a Kantian view, however, both truck drivers, if exhibiting negligence, are equally blameworthy because they both exhibit a bad will, though we may have difficulty *knowing* this, based on the evidence available to us. What actually happens, on Kant's view, is irrelevant. It is worth noting, however, that a subjective consequentialist would also share this "advantage" with the Kantian, or anyone else who took a purely internalist approach to moral evaluation.

4.2 Intuitionistic Approaches

Another body of deontological approaches tries to arrive at substantive principles on the basis of intuitive reactions to cases. There is no theoretical underpinning to these principles aside from the appeal to intuition.[33] One can be either a pluralist (p. 45) or a monist with respect to this approach. The basic idea behind the former is that there may be a variety of factors that go into making an action right or wrong, and it is not possible to reduce these factors to a form that can be presented as a *single*, basic, principle of morality, or a comprehensive moral theory. Thus, no theory such as consequentialism or contractualism grounds the basic principles of morality. This makes the view extremely broad—a

huge variety of principles have been articulated and defended by intuitionist deontologists. It will not be possible to cover all the variation here, but I will look at the most significant.[34]

One might be convinced by a kind of Rossian pluralism that there are a *variety* of different duties, the weight of which cannot be attributed to factors like production of consequences (Ross 1988). One category discussed by Ross and pursued by contemporary writers is that of our special obligations or duties. Indeed, the existence of these duties—i.e. duties to one's family and friends that seem more weighty than duties to strangers—has provided the fodder for much criticism of both utilitarian and Kantian ethics, which advocate impartiality. One needn't be a Rossian pluralist to see the force of these considerations. We do seem to owe more to the near and dear than to others.

Again, one can be a pluralist or a monist. However the intuitionism goes, though, the position is often understood in contrast to consequentialism, in terms of a commitment to some principle(s) with justificatory force independent of an exclusive consideration of consequences.

4.3 Doing and Allowing

Deontologists will frequently hold that one factor with intrinsic moral weight is *doing* harm to others. Doing harm is thought to be much worse than simply allowing harm, even though the effects of both may be completely the same. Further, against consequentialism, there will be situations in which it is wrong to harm someone even if that harm would result in greater good. Suppose that "greater good" could be achieved by killing one innocent person. For example, Alan is a wealthy miser whose wealth, were he to die before squandering it, would be distributed in such a way as to improve the lives of many other people. It would not be permissible, and certainly not obligatory, to kill Alan to ensure the timely redistribution of his assets.[35] Thus, deontological critics of utilitarianism will point to the fact that there must be an 'agent-centred (p. 46) constraint' (I will just use plain 'constraint') against harming in some circumstances.[36] Indeed, some writers characterize deontological approaches in contrast to consequentialist approaches in terms of this one factor. This is a mistake, since the objective consequentialist can argue for constraints on consequentialist grounds.[37] Further, because of the demandingness problem standardly levelled against consequentialism, many deontologists recognize a permission or an option to fail to maximize the good. One need not, for example, give the bulk of one's assets to charity. Such an action would be permitted and even supererogatory, but not required.

Some have noted that there is a morally significant distinction between doing harm and allowing harm. There are circumstances in which agents may permissibly allow harm, though they may not do it. We can refer to this as the doctrine of doing and allowing (DDA).

Philippa Foot has argued for a version of DDA, and against the doctrine of double effect (next section). Her view (Foot 1978) is that there is a morally significant distinction or difference between '(i) initiating or sustaining a harmful causal sequence' and '(ii) (a) allowing or enabling a harmful causal sequence to run its course (b) diverting a harmful causal sequence'.

So, for example, one would not be permitted to run one's car over one person even if doing so were necessary to get to five people in need of rescue. Killing (doing harm) is worse than letting die (allowing harm). Foot would like to offer an account that could accommodate our intuitions regarding two central cases, which were later expanded on by Judith Jarvis Thomson and together constitute 'the trolley problem'. To use Thomson's cases, consider the case of David, a great transplant surgeon, who could save the lives of five patients by sacrificing an innocent person—basically, by taking the life of one person and using his body parts to save the lives of the five. This seems clearly impermissible. But compare this scenario with the case of Edward, the driver of a trolley. The trolley's brakes have failed, and he sees ahead of him on the track five people who are unable to get off the track in time to avoid being hit by the trolley. However, Edward has enough time to turn the trolley onto another track, though, sadly, there is one person on that track who will also not be able to get off in time. Most think that it is morally permissible for Edward to turn the trolley onto the other track, even though that means he will be killing one to save the five. Thus, we have cases that seem to be structurally similar yet intuitively very dissimilar (Thomson 1976). For Foot there are negative and positive duties, negative duties being duties to avoid harm and positive ones being duties to provide aid. The former, she argues, are weightier than the latter. The trolley case differs from the transplant case (or similar cases) in that, in that case, two negative duties conflict; in the transplant case, in contrast, there is a negative duty in conflict with positive duties, and the negative duty trumps.

(p. 47)

Thomson modifies the trolley case by according the agency not to the trolley driver but to a bystander. She then argues that Foot's analysis can't handle the case, since the bystander would be killing one to save five. She contrasts this with the case of a fat man who is pushed onto the trolley tracks to keep it from killing the five. Clearly, she believes, the latter act is immoral; although it involves killing one to save five, it does so in a way that uses another person as a means to saving the five. What makes this action immoral is that the agent is doing something to a person in order to redistribute a threat: 'what matters in these cases in which a threat is to be distributed is whether the agent distributes it by doing something to it, or whether he distributes it by doing something to a person' (Thomson 1976: 16).

Warren Quinn is another example of a philosopher who has argued in favour of the distinction between doing and allowing, against consequentialism (Quinn 1993). Quinn uses several thought experiments to motivate the view that it may sometimes be permissible to allow a harm to befall another person, though not, under the same circumstances, active-

ly to bring about that harm. Quinn's analysis hinges upon noting that there's a difference between positive and negative rights. Negative rights are the rights we have against being harmed by others; positive rights are the rights to actual aid from others. He agrees with Foot's view that negative rights are ones that are more difficult to override than positive ones, but explains this somewhat differently. He argues, against a consequentialist view, that negative rights are more important because respect for negative rights is crucial to our sense of self and moral autonomy:

> A person is constituted by his body and mind. They are parts or aspects of him. For that very reason, it is fitting that he have primary say over what may be done to them—not because such an arrangement best promotes overall human welfare, but because any arrangement that denied him that say would be a grave indignity.... if your life is yours then there must be decisions concerning it that are yours to make—decisions protected by negative rights. (Quinn 1993: 309)

James Rachels, however, has attacked the distinction. He has noted (Rachels 1975) that we can come up with cases in which, holding everything constant but the killing versus letting die aspect, our intuitions are such that we do not hold the killing to be worse than the letting die. What seems to be at the basis of our intuitions about such cases has to do with the agent's intentions; malevolence is worse than indifference.

4.4 Intending Harm

Because of these and other problems, some have argued that the distinction between doing and allowing is not morally significant. Rather, the real distinction is between intending harm and merely foreseeing harm. The constraint, then, is (p. 48) best thought of as a constraint against intending to harm. The doctrine of double effect (DDE) captures this distinction. The DDE holds that it may sometimes be permissible to bring about a harm that is simply foreseen—as a side-effect of some action that is intended to bring about some good end—though it would be wrong to intend that harm as a means to the very same good end. Thus, there is a morally significant distinction between intention and foresight. The classic case summoned to motivate this doctrine is that of a bomber pilot during wartime who has no intention of killing innocent civilians when he drops a bomb on a military target, though he can certainly foresee that some innocent civilians will be killed as an unintended side-effect of the bombing. This may be perfectly permissible, given certain constraints. For example, the bad effect that is merely foreseen must not be out of proportion to the good effect that is intended, and the action itself must not be inherently wrong. However, it would not be permissible for the pilot intentionally to target innocent civilians.

But numerous problems arise. For example, it can prove difficult to maintain a sharp distinction between intention and foresight. To use a case from Alison McIntyre:

> suppose that I divert the trolley off a track with a rare wildflower about to bloom on it and onto a track with one track workman on it, because I judge that perpetuating a nearly extinct species is a more important goal. Would we say that the

harm to the workman is a merely foreseen side effect of saving the wildflower, something wholly unintended? Or would it seem right to say that I intended that the harm fall on him, because I value preserving the flower more? (McIntyre 2001: 236)

McIntyre thinks that a wholesale rejection of intention–foresight would be mistaken. However, her view is that it is impossible to isolate just one way in which the intention–foresight distinction matters morally, and that the various attempts to do so in the literature simply lead to confusion.

As I mentioned earlier, a large number of deontological principles have been suggested as offering a way to navigate these sorts of cases. F. M. Kamm, for example, argues for the principle of permissible harm:

> The basic idea is that it is permissible for (i) greater good and (ii) means that have greater good as their noncausal flip side to cause lesser evil, but not permissible to (iii) intend lesser evil as a means to greater good or to (iv) intend means that cause lesser evil as a foreseen side effect and have greater good as a mere causal effect unmediated by (ii). By 'noncausal flip side,' I mean that the greater good occurring is, in essence, another way of describing the situation in which the means occur. (Kamm 2000: 213–14)

Kamm's idea is that this deontological principle can avoid certain problems. For example, it would sometimes permit harming some to help others: 'Suppose by directing gas into a room we can save five people. However, their breathing normally— the greater good—alters the air flow in the room, redirecting germs, killing an innocent person. In this case it is permissible to use the gas to save five people, because it is the greater good itself which causes the death' (2000: 214). The person (p. 49) who dies in this scenario is not being used as a means. Persons are inviolable on Kamm's approach.

Recently, these approaches have come under fire from moral psychologists co-opting work in psychology. For example, there is compelling evidence that moral intuitions in specific cases are heavily influenced by framing effects, and that various heuristics may have evolved to influence our decision-making by making it adaptive but might have little to offer in terms of reflecting some deep moral truth.[38] The trolley versus transplant cases help to illustrate this finding. Our intuitions about cases that are structurally similar are quite at odds. While there may be a way to resolve the conflict in this particular instance, numerous other cases can be concocted to show that our intuitions are often confused and inconsistent. Thus, simple reliance on intuitionistic responses to cases can be challenged.[39] Authors working within an explicitly theoretical framework have the resources at least to provide an error theory for some intuitions.

The methodology used has also been criticized. Shelly Kagan, for example, points out that the contrast strategy employed by these authors assumes a ubiquity thesis that seems unfounded. Instead of trying, intuitively, to isolate particular morally relevant factors, we are better off relying on foundational moral theory and putting our efforts into developing that alternative (Kagan 1988).[40] As the proposed deontological principles become ever

more baroque in order to respond to cases, there is also the suspicion that the principles merely describe our intuitions and have no prescriptive force.

Thus, a variety of other authors find some theoretical approach, whether consequentialism, Kantian ethics, contractualism, or virtue ethics, to provide a more satisfactory systematic account of moral phenomena.

(p. 50) 4.5 Contractarian and Contractualist Ethics

These approaches could be considered deontological since they also treat the right as prior to the good. The contractarian approach has generally been concerned with issues in political philosophy, though some writers have felt that it could be extended more broadly to cover norms regulating individual and personal conduct as well. Fairly recently, a distinction has been drawn between contractarian and contractualist ethics.

The contractarian approach proceeds with the acknowledgement that normative ethics guides action in a social world. This purpose is achieved by acceptance on the part of rational persons within a given society of reasonable rules to guide action. Thus, the justification for the rules is a kind of contract or agreement between rational people. The contract is not based on the view that morality aims at a particular good, and thus is not teleological. This agreement is generally not regarded as "actual"; rather, the principles governing right action are those that would be agreed upon by rational self-interested persons through a process of negotiation. If the agreement needed to be historical fact and actual agreement, then it would not bind currently existing people: the obligations incurred by our ancestors are not ours. Thus, the contract is generally understood as "hypothetical"—what rational persons would agree to under certain circumstances. But the notion of a binding hypothetical contract has been attacked by Dworkin, who argues:

> Suppose that I did not know the value of my painting on Monday; if you had offered me $100 for it then I would have accepted. On Tuesday I discovered it was valuable. You cannot argue that it would be fair for the courts to make me sell it to you for $100 on Wednesday. It may be my good fortune that you did not ask me on Monday, but that does not justify coercion against me later. (Dworkin 1977: 152)

One contractarian, David Gauthier (1986), follows Hobbes. Gauthier takes a central feature of morality to be the issue of *justified constraint*. Contractarianism gives an account of this feature in terms of what rational agents would agree to. The just moral constraints on behaviour are those that rational persons would agree to in recognition of the fact that they interact with other agents. Gauthier tries to bypass Dworkin's concern by noting that appealing to a contract is a way of providing a thought experiment that gets us to recognize what rationality demands. For Gauthier, the bargainers are characterized as self-interested. Thus, the consent to abide by moral principles and the content of those principles is driven by self-interest. So for Gauthier the contract is a bargain in which each of the bargainers tries to negotiate the rules that best achieve his or her utility. However, he is willing to make a concession to solve the "compliance problem". To solve this problem he argues in favour of a principle of "minimax concession" (or "maximin relative benefit") according to which the ratio-

nal bargainer will be concerned to minimize the (p. 51) maximum relative concessions she makes in the bargaining process.[41] The bargainer must perceive her concessions to be reasonable, and this is judged relative to what others are willing to give up or concede. She needs to perceive that she is optimizing the outcome for herself in such a way as to secure compliance from others. A major difficulty of this approach, however, is that in order to guarantee compliance after the contract is adopted, Gauthier argues that rational individuals will make a choice to be constrained maximizers of self-interest. He believes that they will choose to cultivate dispositions to keep their agreements, etc., and leave to one side self-interest in situations where they perceive that others are so acting. This can lead to situations in which the commitment to do what one says one will do can have apocalyptic results. Gauthier's most recent position tries to avoid this problem by holding that it is rational to act on a prior intention if and only if the expected outcome is as good as the expected outcome of not having formed that intention. But Joe Mintoff (1997) points out that while this formulation gets Gauthier out of the problem of apocalyptic retaliation strategies, it leaves him with other problems. Consider a variation on the famous toxin puzzle, presented by Gregory Kavka (1983). In this case, you are to imagine that an eccentric tycoon presents you with a vial of toxin and tells you that if you drink it you will not be caused permanent damage, but you will experience one day of intense illness. However, he also tells you that if at midnight tonight you form the intention to drink the toxin at eight tomorrow morning, he will pay you $1 million. You need not actually drink the toxin to get the money; you need only form the intention to do so. If you change your mind after forming the intention, you still get the money. It seems that on Gauthier's account it is rational for you to go ahead nevertheless and drink the toxin.[42]

Again, on the contractarian approach the right action will be specified according to rules of cooperation between reasonable persons. The spirit of the approach is to determine what rules would be in everyone's interests. Which principles bind us is a matter of agreement or contract. Theoretically, this would be a matter of negotiation.

And this is where contractarians and contractualists part company. One problem is that merely negotiated principles intuitively seem to have no binding moral force. Negotiation is not a legitimate basis for *moral* principles because those principles would then simply be arbitrary: Suppose it was agreed that the weakest members of a community be killed. Again, just intuitively, it seems that this can't be right; and, *even if* it were right, it's not right on the basis of an agreement. Contractualists do hold that moral principles are agreed upon, but they are inspired in making this claim by Kant's notion of equal respect between persons. (p. 52) Principles agreed upon merely as the result of bargaining have no binding force. The legitimate principles are those that would be agreed to from the point of view of 'free and equal persons' operating under conditions of mutual respect. Rawls's methodology in *A Theory of Justice* embraces this Kantian approach. Though Rawls is limiting the scope of his work to justice, and is not offering a comprehensive theory of morality, one could see how the methodology could be employed to that end.[43] The correct principles of justice are those agreed upon by free and equal persons under conditions that ensure fairness of choice—that is, behind a "veil of ignorance" in what Rawls dubs the "original position" since it refers to the initial bargaining point. The agreement

is hypothetical. He argues that rational agents will agree to two principles of justice under these conditions (Rawls 1971). The first principle is 'Each person is to have an equal right to the most extensive total system of equal basic liberties compatible with a similar system of liberty for all.' The second principle is more complicated, but basically holds that inequalities are to be arranged so that they are 'to the greatest benefit of the least advantaged…and… attached to offices and positions open to all under conditions of fair equality of opportunity'. Numerous criticisms have been made of this approach, in terms of both the methodology and the content of some of Rawls's assumptions.

One such criticism is that Rawls is not really offering a contractualist approach at all—that the method of reflective equilibrium is sufficient to generate the intuitions Rawls is counting on in support of his principles of justice. Recall Dworkin's criticism of the hypothetical contract as a basis for obligation. Because of this criticism Dworkin believes that Rawls's argument is best understood as simply drawing our attention to some other argument for the two principles of justice. Rawls seems to acknowledge this when he writes that his argument is one that seeks 'to draw solely upon basic intuitive ideas that are embedded in the political institutions of a constitutional democratic regime and the public traditions of their interpretation'(Rawls 1985: 225). Thus, we don't need a bargaining situation at all— just one rational individual reflecting on what would be reasonable.[44] But the notion that Rawls employs here also comes under attack. Rawls believes that the maximin principle is the rational principle to employ in the original position. Critics have noted that this is to assume that human beings are risk-averse, or that we ought to be risk-averse if we are rational (at least, in certain narrowly specified contexts).

One line of attack is to point out that the streamlined individual in Rawls's original position, behind the veil of ignorance, is no individual at all—disembodied and sexless.[45] This general line of attack is also developed by communitarian critics. For example, Michael Sandel (1982) argues that Rawls's conception of the self (p. 53) doesn't reflect the reality of how selves are formed by their goals, which are discovered by individuals within a shared community. Sandel writes: 'the Rawlsian self is not only a subject of possession, but an antecedently individuated subject, standing always at a certain distance from the interests it has. One consequence of this distance is to put the self beyond the reach of experience, to make it invulnerable, to fix its identity once and for all' (1982: 62).

This means, on Sandel's view, that Rawls's account rules out views of the self not committed to some 'antecedently individuated self'.

Be that as it may, Rawls's invocation of Kantian elements was responsible for generating a novel approach within this tradition—one currently labelled contractualism.

Central to some articulations of contractualism is the view that agents are accountable to others. Thus, these principles must be ones that can be justified to others. They must be ones that others could not reasonably reject. As Scanlon puts it: 'An act is wrong if its performance under the circumstances would be disallowed by any system of rules for the

general regulation of behavior which no one could reasonably reject as a basis for informed, unforced, general agreement' (Scanlon 1982: 153).

And, later (1998): 'an act is wrong if it would be disallowed by any *principle* that no one could reasonably reject'. Note that on this view, as opposed to the consequentialist view, well-being is not given a central role (Scanlon 1998: 215 ff.). It is not the case that well-being is irrelevant—of course consideration of well-being will figure into the reasonable rejection of principles—it is just that well-being is not foundational to this account in the way it is to consequentialism. This avoids problems, Scanlon maintains, that a consequentialist will have through privileging well-being over other significant factors.

While the contractualist approach avoids a serious intuitive difficulty associated with the contractarian approach, they both share the view that agreement is crucial to determining moral principles and rules, and, thus, to giving an account of right action. This overall approach is hampered by an inclusion problem. Some writers have pointed out that contractarianism–contractualism has a general problem with including the vulnerable. Persons who are not in a position to negotiate or engage in rational reflection because they lack the requisite capacities are left out.[46] This problem applies to animals as well. Scanlon tries to address this problem by distinguishing broad and narrow conceptions of morality: contractualism is only a narrow conception that deals specifically with the norms governing interactions between rational beings. This is the relevant scope of his account. Animals fall outside that scope, but that is not to say they have no moral standing whatsoever. However, that would require a different argument. Scanlon's theory is at best incomplete. The problem that most critics focus on is its redundancy. The basic (p. 54) idea is that the appeal that Scanlon makes to what no one could 'reasonably reject' makes his account redundant, since we might just as well make direct reference to the grounds of reasonable rejection for our criterion or criteria for wrongness.[47] Thus, the appeal to what would be reasonably rejected does no work. In the end, we have a direct appeal to moral value that actually does the work.

These approaches have also been the target of many attacks from feminist philosophers and virtue ethicists, since the inclusion problem has historically been associated with the exclusion of women, for example, from the agreement. Of course, exclusion of groups of persons could not be endorsed by the theory: any contractualist is committed to including all rational persons. This may be more of a problem for the contractarian, since there seems to be no principled way for arguing that the scope of the agreement *must* include all persons.[48] Importing a moral notion of fairness would be to acknowledge such norms independently of the agreement itself.

Still, in recent years one sees many feminist writers moving away from wholesale criticism of contractarian ethics, and, indeed, embracing a variety of contractarianism that is more amenable to feminist concerns. For example, Susan Moller Okin extends the Rawlsian project to include the family as a social institution. Okin revises the procedure in the original position by advocating a kind of empathic regard on the part of the agents in the original bargaining position: 'Rawls' theory of justice is itself centrally dependent upon

the capacity of moral persons to be concerned about and to demonstrate care for others, especially others who are most different from themselves' (Okin 1989: 247). Jean Hampton (1993) offers a defence of the contractarian approach, arguing one can be a feminist and a contractarian: justice is a value that operates between friends and family members, as well as members of civil society. Contractarian ethics can speak to these concerns, even if the tradition might seem to have left them out.

5. Virtue Ethics

Anscombe inspired widespread disaffection with standard ethical theories with her very influential article 'Modern Moral Philosophy' (1981). There she argued that abstract normative concepts such as "ought" or "moral obligation" make no sense outside a legislative context. Thus, given the absence of God in most accounts of (p. 55) what we morally ought to do, she holds that modern moral philosophy should just give it up and instead look to psychology for guidance. There should be greater focus on moral psychology and the virtues. Taking this seriously, many philosophers have come to favour a virtue-ethical approach to moral evaluation. After some disagreement about how exactly to define "virtue ethics", the field seems to have centred on the following claim: virtue ethics treats virtue evaluation as the "primary" mode of evaluation; thus, any account of right action that is virtue-ethical needs to define right action in terms of virtue (otherwise, the account of virtue is derivative of the account of right action).

Several writers have attempted to do this, most notably Rosalind Hursthouse and Michael Slote. Hursthouse defines "right action" in the following way:

> An action is right iff it is what a virtuous agent would, characteristically, do in the circumstances, except for tragic dilemmas, in which a decision is right iff it is what such an agent would decide, but the action decided upon may be too terrible to be called 'right' or 'good.' (And a tragic dilemma is one from which a virtuous agent cannot emerge with her life unmarred.) (Hursthouse 1999: 79)

Leaving aside the issue of tragic dilemmas, the idea that Hursthouse is appealing to is that we use the virtuous agent as an exemplar of sorts.[49] However, her procedure is underdescribed in that it is not clear what circumstances we should consider in applying it. For example, suppose that Bill knows that he buckles under pressure— he is less than virtuous. He then finds himself in a situation in which some people in a burning building need help, and he is in a position to help them. The "virtuous" person would run into the burning building (assuming that the objective risk to the person is small). But Bill knows that once inside the building, confronted with flames and smoke, he would freeze and be unable to help. Should he then run into the burning building? If Hursthouse's procedure leads to an affirmative response, it is mistaken.[50] But if we are to build in imperfections, and ask what a person who was not virtuous would do, then the account appears incoherent.

Hursthouse's approach is neo-Aristotelian. Michael Slote has recently developed an account with Humean roots—one that advocates a form of sentimentalism.

Slote's account is distinctive in that it is a pure form of evaluational externalism. On Slote's view, the moral quality of an action depends on the agent's motives in that it must be an expression of the agent's virtuous motives. Thus, the standard of "rightness" will be a purely internal one. What is brought about by the action is irrelevant. Slote has a problem, which he is aware of, in that such an account can be charged with moral autism, or being too insular and self-absorbed. He counters this by claiming that a person who has a good motivational structure will care (p. 56) about what happens in the world; thus, the good agent is not entirely cut off from what actually happens. This manoeuvre has some success; however, it leaves open the possibility of a person performing a right action and exemplifying virtuous motivation, even though what he does is always a disaster. Such a moral klutz would nevertheless be acting rightly as long as he cared about others and tried to make his actions effective in the proper way. Actual failure to make his actions effective in this way does not affect the rightness of those actions, on Slote's account. The advantage of such an internal approach is that it provides great insulation against moral luck. As with Kant, what matters to the moral quality of what you do is your internal psychological states. What happens in the world is subject to all kinds of luck factors over which the agent ultimately has no control. And without control, there can be no responsibility. That being said, this approach seems problematic in that many persons, even those who are not consequentialists, will hold that consequences matter morally—it is just that they may not be the only thing that matters. Slote abandons this.

An additional worry has to do with how one can spell out what "expresses" a virtue without circularity. Slote has to argue that the right action is an expression of virtue and not simply caused by the virtue. This is because there may be particular occasions on which a benevolent person fails to actually act benevolently (Slote 2001: 17). So, the act must express the benevolence of the person in order to be the right act. Slote also needs to add this condition to rule out cases of neutral actions that might be caused by benevolent motives: it may be that benevolent persons eat red apples instead of green ones, but this doesn't count since eating a red apple is not an expression of benevolence. However, this puts a real burden on Slote to explain what exactly is meant by "expresses" benevolence. If he does this by noting that these actions must be benevolent ones, the circularity worry comes up.[51]

Another worry with internalist approaches in general is that, although they score points for insulating agents against moral luck, they also seem to cut the agent off from the world. Hursthouse's brand of virtue ethics avoids this problem since the virtues, while characterized by certain psychological states of the agent, must also somehow contribute to human flourishing, and in her case she ties the flourishing to that of the virtuous agent. Evidence for this, she thinks, is that parents tend to educate their children to be virtuous. And, presumably, parents are very concerned with the well-being of their children. Thus, there is some evidence that we hold the virtues to contribute to individual flourishing. Some will find Hursthouse's approach a bit more intuitively plausible in this regard than Slote's, though in my opinion Hursthouse's approach still has the problem of placing too many psychological requirements on virtue. Though she softens this requirement in her account, her account would still rule out what I have elsewhere (Driver 2001, ch. 2)

(p. 57) called 'the virtues of ignorance'—traits like modesty that seem, at least in some cases, to involve actual ignorance on the part of the virtuous agent.

6. Feminist Ethics

Another strand has been inspired by the work of Carol Gilligan. Her book *In a Different Voice* (1982) raised the issue of gender differences in moral thinking. By attacking Lawrence Kohlberg's account of moral development in children, Gilligan pointed to the possibility of significant gender differences in ethical thought. This led, or fed into, at any rate, the challenge to "justice" or "rule"-based systems of morality. Female experience was said to be more particularistic, caring, and nurturing, commensurate with the sort of experiences most females have and the norms that tend to govern their relationships with others. Nel Noddings (1984) developed an "ethics of care" approach that incorporated these particularistic insights and portrayed the "justice" approach as rigid and cold, neglecting the emotional. Indeed, this has been a hallmark of early feminist ethics—the claim that a very masculine sense of "reason" and "rationality" has dominated ethical evaluation, at the expense and even denigration of emotion. The 'ethics of care' approach has developed its own following, particularly in health care ethics. Again, there is a connection to work being done in virtue ethics. Michael Slote views his virtue-ethical approach as one that develops an ethics of care that is responsive to many of the criticisms of standard ethics raised by feminist philosophers.

However, as Samantha Brennan has pointed out (1999), feminist ethics has come a long way from its early roots. Indeed, both Gilligan's work and Nel Noddings's development of the ethics of care have come under increasing critical scrutiny among feminist philosophers. For example, Alison Jaggar (1991) has suggested that, though the care approach has been an important development, particularly because of its criticism of the standard theories, it is also flawed because it is too narrow an approach to ethics. Elizabeth Spelman (1991) has pointed to the dangers of simply focusing on a contrast between "feminine" care and "masculine" justice, when the reality is more complex.[52]

Additionally, earlier feminist philosophers had suggested a radical approach to ethical theory. Since traditional ethical theory, and, it was thought, associated (p. 58) concepts such as "rights", were skewed by only accounting for male experience, such concepts should be jettisoned. This claim is reminiscent of the sort of eliminativism one saw in virtue ethics, although the reasons for it are different. However, this eliminativism has been rejected by other feminist ethicists who, instead, try to develop feminist accounts of "rights" and "responsibilities". Rather than attacking the concepts themselves, they focus their attack on non-feminist understandings of the concepts, which are taken to be defective because of their individualism and their failure to incorporate relational factors. The challenge to this relational approach, however, is to avoid extreme relativism. While some would embrace relativism, many would not, since it would undercut the revisionist aspirations of many feminist ethical theorists. The relational approach seems inconsistent with the view that certain practices elsewhere are just plain wrong—practices that involve the

subjection of women and other vulnerable groups. This is a serious problem for an approach motivated by a desire to see reforms implemented that would contribute to the equality and well-being of women. Along these lines, Alison Jaggar writes:

> Against the relativist tendency in feminist ethics should be set feminism's concern that its moral critique of the practices (and theory) of the larger society—and perhaps even the practices (and theory) of other societies—should be objectively justified. Feminist ethics recognizes that we inhabit a painfully prefeminist world and takes itself to be contributing to the transformation of this world into one in which the basic moral commitments of feminism have become universally accepted. (Jaggar 1991: 94–5)

Feminist conceptions of moral theory cover a very wide range of different views with varying theoretical commitments. They are joined, however, in the belief that standard ethical theories have undervalued women's experience and concerns.

7. Conclusion

Normative ethical theory has undergone a transformation in the last generation. Challenges have been made to normative ethical theory—particularly to the commitment to impartiality and the view that there is a single moral principle sufficient to guide action. Greater focus on relationships, virtues, and less abstract issues has transformed the development of the major theories. Consequentialism, in particular, has shown great resilience in meeting the challenges posed by critics who are concerned about placing excessive weight on impartiality and oversimplification. However, a future challenge remains. Research in psychology may yield explanations for some of our intuitions that reveal *explanations* that call into question the justifying force of intuition. To the extent that all the theories we have (p. 59) considered here have, to some degree, felt compelled to conform to commonsense intuitions, they may be mistaken. It may well be that normative ethical theory should embrace reformist goals.

References

Adams, Robert (1976). 'Motive Utilitarianism'. *Journal of Philosophy*, 73: 467–81.

Anscombe, G.E.M.(1981). 'Modern Moral Philosophy', in *The Collected Papers of G. E. M. Anscombe: Ethics, Religion and Politics*. Minneapolis: University of Minnesota Press.

Audi, Robert, and Sinnott-Armstrong, Walter (eds.) (2002). *Rationality, Rules, and Ideals*. Lanham, Md.: Rowman & Littlefield.

Bales, Eugene (1971). 'Act-Utilitarianism: Account of Right-Making Characteristics or Decision-Making Procedure?' *American Philosophical Quarterly*, 8: 257–65.

Baron, Marcia (1995). *Kantian Ethics Almost Without Apology*. Ithaca, NY: Cornell University Press.

Bentham, Jeremy (1823). *Introduction to the Principles of Morals and Legislation*. London: W. Pickering.

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

Brennan, Samantha (1999). 'Recent Work in Feminist Ethics'. *Ethics*, 109: 858–93.

Brink, David (1989). *Moral Realism and the Foundations of Ethics*. Cambridge: Cambridge University Press.

Broome, John (1991). *Weighing Goods*. Oxford: Blackwell.

Carritt, E. F. (1947). *Ethical and Political Thinking*. Oxford: Clarendon Press.

Cummiskey, David (1996). *Kantian Consequentialism*. New York: Oxford University Press.

Dreier, James (ed.) (forthcoming). *Contemporary Debates Ethical Theory*. Malden, Mass.: Blackwell.

Driver, Julia (1992). 'Caesar's Wife: On the Moral Significance of Appearing Good'. *Journal of Philosophy*, 89: 331–43.

——(2001). *Uneasy Virtue*. New York: Cambridge University Press.

——(2003). Review of Michael Slote, *Morals from Motives. Journal of Ethics*, 7: 233–7.

—— (forthcoming). 'Virtue Theory', in James Dreier (ed.), *Contemporary Debates in Ethical Theory*. Malden, Mass.: Blackwell.

Dworkin, Ronald (1977). *Taking Rights Seriously*. Cambridge, Mass.: Harvard University Press.

Feldman, Fred (1997). *Utilitarianism, Hedonism, and Desert: Essays in Moral Philosophy*. Cambridge: Cambridge University Press.

Finnis, John (1980). *Natural Law and Natural Rights*. Oxford: Clarendon Press.

Foot, Philippa (1978). 'The Problem of Abortion and the Doctrine of Double Effect', in Foot, *Virtues and Vices and Other Essays*. Berkeley: University of California Press.

Gauthier, David (1986). *Morals by Agreement*. New York: Clarendon Press.

——(1990). *Moral Dealing*. Ithaca, NY: Cornell University Press.

George, Robert (1993). *Making Men Moral*. Oxford: Clarendon Press.

Gert, Bernard (1998). *Morality*. New York: Oxford University Press.

Gilligan, Carol (1982). *In a Different Voice*. Cambridge, Mass.: Harvard University Press.

Hampton, Jean (1993). 'Feminist Contractarianism', in Louise Antony and Charlotte Witt (eds.), *A Mind of One's Own*. Boulder, Colo.: Westview Press.

(p. 60) Harman, Gilbert (1977). *The Nature of Morality*. New York: Oxford University Press.

Herman, Barbara (1993). *The Practice of Moral Judgment*. Cambridge, Mass.: Harvard University Press.

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

Horowitz, Tamara (1998). 'Philosophical Intuitions and Psychological Theory'. *Ethics*, 108: 367–85.

Hurka, Thomas (1993). *Perfectionism*. New York: Oxford University Press.

Hursthouse, Rosalind (1999). *On Virtue Ethics*. New York: Oxford University Press.

Jackson, Frank, and Pargetter, Robert (1986). 'Oughts, Options, and Actualism'. *Philosophical Review*, 95: 233–55.

Jaggar, Alison (1991). 'Feminist Ethics: Projects, Problems, Prospects', in Claudia Card (ed.), *Feminist Ethics*. Lawrence: University Press of Kansas.

Kagan, Shelley (1988). 'The Additive Fallacy'. *Ethics*, 99: 5–31.

—— (1989). *The Limits of Morality*. New York: Oxford University Press.

——(1998). *Normative Ethics*. Boulder, Colo: Westview Press.

Kahneman, Daniel, and Tversky, Amos (1979). 'Prospect Theory: An Analysis of Decision Under Risk'. *Econometrica*, 47: 263–91.

Kamm, Frances (1998). 'Moral Intuitions, Cognitive Psychology, and the Harming-versus-Not-Aiding Distinction'. *Ethics*, 108: 463–88.

—— (2000). 'Nonconsequentialism', in *The Blackwell Guide to Ethical Theory*. Malden, Mass.: Blackwell.

Kant, Immanuel (1969). *Foundations of the Metaphysics of Morals*, trans. Lewis White Beck, with critical essays ed. Robert Paul Wolff. New York: Macmillan.

Kavka, Gregory (1983). 'The Toxin Puzzle'. *Analysis*, 43: 33–6.

Kittay, Eva (1999). *Love's Labor: Essays on Women, Equality, and Dependency*. London: Routledge.

Korsgaard, Christine (1983). 'Two Distinctions in Goodness'. *Philosophical Review*, 2: 169–95.

—— (1986). 'The Right to Lie: Kant on Dealing with Evil'. *Philosophy and Public Affairs*, 15: 325–49.

—— (1996). 'From Duty and for the Sake of the Noble: Kant and Aristotle on the Morally Good Action', in Stephen Engstrom and Jennifer Whiting (eds.), *Aristotle, Kant, and the Stoics*. Cambridge: Cambridge University Press.

Lockhart, Ted (2000). *Moral Uncertainty and its Consequences*. New York: Oxford University Press.

Lyons, David (1965). *The Forms and Limits of Utilitarianism*. Oxford: Oxford University Press.

McCloskey, H.J. (1957).'"Two Concepts of Rules"—a Note'. *Philosophical Review*, 66: 466–85.

McGinn, Colin (1999). 'Reasons and Unreasons'. *New Republic*, 220: 34–8.

McIntyre, Alison (2001). 'Doing Away with Double Effect'. *Ethics*, 111: 219–55.

Malm, Heidi (1992). 'In Defense of the Contrast Strategy', in John Martin Fischer and Mark Ravizza (eds.), *Ethics: Problems and Principles*. Fort Worth, Tex.: Harcourt, Brace, Jovanovich.

Mintoff, Joe (1997).'Rational Cooperation, Intention, and Reconsideration'. *Ethics*, 107: 612–43.

Moore, G. E. (1988). *Principia Ethica*. Amherst, NY: Prometheus Books.

Morris, Christopher (1991). 'Moral Standing and Rational-Choice Contractarianism', in Peter Vallentyne (ed.), *Contractarianism and Rational Choice*. New York: Cambridge University Press.

Mulgan, Tim (2001). *The Demands of Consequentialism*. Oxford: Clarendon Press.

(p. 61) Murphy, Liam (2000). *Moral Demands in Nonideal Theory*. New York: Oxford University Press.

Nagel, Thomas (1979). 'The Fragmentation of Value', in Nagel, *Mortal Questions*. New York: Cambridge University Press.

Noddings, Nel (1984). *Care: A Feminine Approach to Ethics and Moral Education*. Berkeley: University of California Press.

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

Okin, Susan (1989). *Justice, Gender, and the Family*. New York: Basic Books.

O'Neill, Onora (1996). 'Kant's Virtues', in Roger Crisp (ed.), *How Should One Live?* Oxford: Clarendon Press.

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

Pateman, Carole (1988). *The Sexual Contract*. Stanford, Calif.: Stanford University Press.

Pettit, Philip (1984). 'Satisficing Consequentialism', p. II. *Proceedings of the Aristotelian Society*, suppl. vol., 58: 165–76.

—— (1993). *The Common Mind*. New York: Oxford University Press.

—— (1997). 'The Consequentialist Perspective', in Marcia Baron, Philip Pettit, and Michael Slote, *Three Methods of Ethics: A Debate*. Malden, Mass.: Blackwell.

—— (1999). 'Doing Unto Others'. *Times Literary Supplement* (25 June), 7–8.

—— and Smith, Michael (2000). 'Global Consequentialism', in Brad Hooker, Elinor Mason, and Dale E. Miller (eds.), *Morality, Rules, and Consequences*. Lanham, Md.: Rowman & Littlefield.

Quinn, Warren (1993). 'Actions, Intentions, and Consequences: The Doctrine of Doing and Allowing', in Quinn, *Morality and Action*. Cambridge: Cambridge University Press.

Rachels, James (1975). 'Active and Passive Euthanasia'. *New England Journal of Medicine*, 292: 78–80.

Railton, Peter (1984). 'Alienation, Consequentialism, and the Demands of Morality'. *Philosophy and Public Affairs*, 13/2: 134–71. Repr. in Samuel Scheffler (ed.), *Consequentialism and its Critics*. Oxford: Oxford University Press, 1988.

—— (1989). 'Naturalism and Prescriptivity'. *Social Philosophy and Policy*, 7: 151–74.

Rawls, John (1955). 'Two Concepts of Rules'. *Philosophical Review*, 64: 3–13.

——(1971). *A Theory of Justice*. Cambridge, Mass.: Belknap Press.

——(1985). 'Justice as Fairness: Political not Metaphysical'. *Philosophy and Public Affairs*, 14: 223–51.

Richards, David (1971). *A Theory of Reasons for Action*. Oxford: Clarendon Press.

Ridge, Michael (2001). 'Saving Scanlon: Contractualism and Agent-Relativity'. *Journal of Political Philosophy*, 9: 472–81.

Ross, W. D. (1988). *The Right and the Good*. Indianapolis: Hackett.

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

Scanlon, Thomas (1982). 'Contractualism and Utilitarianism', in Amartya Sen and Bernard Williams (eds.), *Utilitarianism and Beyond*. Cambridge: Cambridge University Press.

—— (1998). *What we Owe to Each Other*. Cambridge, Mass.: Harvard University Press.

Scheffler, Samuel (2000). *The Rejection of Consequentialism*. Oxford: Clarendon Press.

Singer, Peter (1972). 'Famine, Affluence, and Morality'. *Philosophy and Public Affairs*, 1: 229–43.

Sinnott-Armstrong, Walter (1988). *Moral Dilemmas*. Oxford: Blackwell.

Skorupski, John (1995). 'Agent-Neutrality, Consequentialism, Utilitarianism…A Terminological Note'. *Utilitas*, 7: 49–54.

(p. 62) Slote, Michael (1984). 'Satisficing Consequentialism', p. I. *Proceedings of the Aristotelian Society*, suppl. vol., 58: 139–63.

—— (2001). *Morals from Motives*. Oxford: Oxford University Press.

Smart, J. J. C., and Williams, Bernard (1973). *Utilitarianism: For and Against*. Cambridge: Cambridge University Press.

Smith, Michael (2001). 'Immodest Consequentialism and Character'. *Utilitas*, 13: 173–94.

Spelman, Elizabeth (1991). 'The Virtue of Feeling and the Feeling of Virtue', in Claudia Card (ed.), *Feminist Ethics*. Lawrence: University Press of Kansas.

Stocker, Michael (1976). 'The Schizophrenia of Modern Ethical Theories'. *Journal of Philosophy*, 73: 453–66.

—— (1990). 'Friendship and Duty: Some Difficult Relations', in Owen Flanagan and Amélie Oksenberg Rorty (eds.), *Identity, Character, and Morality: Essays in Moral Philosophy*. Cambridge, Mass.: MIT Press.

Thomson, Judith Jarvis (1976). 'Killing, Letting Die, and the Trolley Problem'. *The Monist*, 59: 204–17.

van Roojen, Mark (1999). 'Reflective Moral Equilibrium and Psychological Theory'. *Ethics*, 109: 846–57.

Watson, Gary (1998).'Some Considerations in Favor of Contractualism', in Watson, *Rational Commitment and Morality*. Cambridge: Cambridge University Press.

Wolf, Susan (1982). 'Moral Saints'. *Journal of Philosophy*, 79: 419–39.

Notes:

(1) For utilitarianism specifically, Eugene Bales (1971) made this distinction.

(2) See Rawls (1971) for a discussion of this way of drawing the distinction. A teleological theory does

(3) See e.g. the discussion in Sinnott-Armstrong (1988).

(4) Much recent work has been done that attempts to clarify the notion of "intrinsic good" or "intrinsic value" and the contrast with extrinsic and instrumental good or value. Christine Korsgaard, for example, argues that though intrinsic is often contrasted with instrumental, this is mistaken; instead, the contrast is with extrinsic or relational good or value. Instrumental is instead contrasted with final good, or the value a thing has as an end in itself. Intrinsic value is the value something has "in itself", the value it carries with it, so to speak, and is non-relational. See Korsgaard (1983: 169–95). While Korsgaard has pointed out an important confusion, usage is such that I will still use "intrinsic" to refer to final good.

(5) Jeremy Bentham famously lists a host of other parameters for measuring pleasure. See his discussion in Bentham (1823)

(6) Walter Sinnott-Armstrong (1988) makes this point.

(7) John Broome (1991) takes up a discussion of the problems associated with weighing. He also discusses one of the major problems of ethical theory (and not just consequentialist theories), which is how to deal with risk or uncertainty. Space precludes a full discussion of this very important issue, but I refer the reader to Broome's book, particularly the discussion in ch. 6. For a discussion of ethical decision-making and uncertainty, see also Ted Lockhart (2000).

(8) See his essays on hedonism in p. II, 'Hedonism', of Feldman (1997).

(9) A number of people have raised this criticism. For example, see David Brink's discussion (1989).

(10) However, it is worth noting that teleological theories such as utilitarianism don't have a lock on the view that well-being is of central importance to morality. Indeed, Gary Watson argues that contractualism can better account for the intuition that well-being is central to normative ethical theory while accommodating our intuitions about deontological constraints on action. See Watson (1998).

(11) Shelly Kagan makes use of a distinction that, unfortunately, space considerations prevent me from fully utilizing in this chapter. Kagan classifies theories as "teleological" or "deontological" at two different levels: the factoral and the foundational. Given this, it should be noted that the intuitionistic deontologist is deontological at the factoral level (the level of figuring out what to do) but is non-foundational since there is no deeper theoretical level to appeal to in providing justification for the principles used to help us figure out what to do. See Kagan (1998).

(12) Philip Pettit in part characterizes consequentialist approaches, of which utilitarianism is the best known, as theories that hold that rightness is a matter of promoting the good as opposed to honouring the good. See e.g. Pettit (1997).

(13) I believe that it is misleading to characterize consequentialist theories as committed to agent-neutrality. As John Skorupski notes, what a consequentialist does is define rightness in relation to a theory of intrinsic value. But 'a theory of intrinsic value may generate agent-neutral or (as with Sidgwick's rational egoist) agent-relative reasons'. See his discussion in Skorupski (1995).
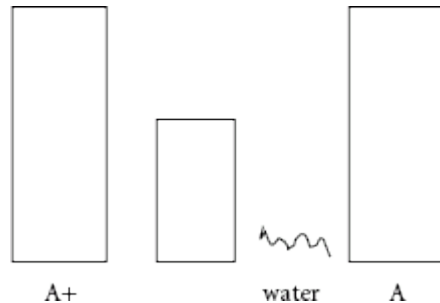
(14) Parfit plausibly argues that lowering average utility by 'mere addition' cannot be bad. He writes: 'There is *Mere Addition* when, in one of two outcomes, there exist extra people (1) who have lives worth living, (2) who affect no one else, and (3) whose existence does not involve social injustice' (1984: 420). If we compare A and A, A seems to be better even though the additional persons lower the average. This is just a part of his presentation of the mere addition paradox. For a discussion of the entire paradox and not merely this claim, see Parfit (1984: 420 ff.).

(15) This *type* of case was presented in Carritt (1947) and discussed by Rawls (1955), and in McCloskey's discussion of Rawls's article (1957).

(16) See also, for his this point, Cummiskey (1996).

(17) A number of writers have argued that impartial ethics makes too many demands in viewing moral reasons as overriding (thus, this isn't simply a problem for utilitarianism). See e.g. Susan Wolf's excellent (1982) for more on this issue.

(18) See Slote (1984) and Philip Pettit's response (1984).

(19) Again, see the discussion in Pettit (1984). See also Philip Pettit and Michael Smith's discussion of global versus local consequentialism in (2000).

(20) For critical discussion of this approach, see Mulgan (2001).

(21) Rule consequentialism will be discussed in more detail in what follows, as will the objective approach. For another indirect approach, see Adams (1976). Philip Pettit and Michael Smith discuss local forms of consequentialism, motive utilitarianism being one example, and contrast this form with global consequentialism. Their view is that global consequentialism is superior in that it does not privilege any one evaluand (e.g. motives, rules, etc.) over another: if consequentialism offers a sound form of evaluation for one, why not the others? See Pettit and Smith (2000).

(22) The latter is discussed by Michael Stocker (1976).

(23) David Lyons, for example, notes that though act and rule utilitarianism may be intensionally non-equivalent, expanding the content of the "rules" to avoid counter-intuitive absolutism will make them extensionally equivalent. Thus, for all practical purposes, rule utilitarianism collapses in act utilitarianism. See his discussion in Lyons (1965). See also J. J. C. Smart's discussion of the collapse problem in Smart and Williams (1973: 3–74).

(24) See Derek Parfit's presentation of this problem in Parfit (1984).

(25) Thus, the account I offer of consequentialism, and the account I believe is supported by Skorupski's discussion of agent-neutrality and consequentialism, is not *necessarily* incompatible with that offered by Michael Smith (2001: 174) when he writes that 'a distinctive mark of a consequentialist theory, as opposed to a non-consequentialist theory, is that it takes moral values to be neutral, as opposed to relative'. On the objective approach discussed here, agent-relative factors are justified by appeal to neutral reasons: these factors help to promote overall well-being.

(26) See Skorupski (1995).

(27) Gert views himself as advocating a natural law approach in which the moral rules are discovered by human reason in the sense that they are the rules all rational persons accept; See Gert (1998). For critical discussion of Gert's claims, see Audi and Sinnott-Armstrong (2002).

(28) For example, Barbara Herman is uncomfortable with this classification since it seems to hold that Kant's ethics does not have a firm grounding in value—a claim with which she strongly disagrees. However, she does agree that in some of the weaker senses of "deontological"—i.e. as simply being opposed to maximizing the good, or as being committed to the moral significance of certain distinctions independent of a consideration of consequences—Kant's theory is "deontological". See Herman (1993: 208–40).

(29) Though not a virtue ethicist, Onora O'Neill covers this issue in O'Neill (1996).

(30) See e.g. Herman (1993: 1–22).

(31) See e.g. Korsgaard (1996).

(32) I discuss this issue more in Driver (1992).

(33) Of course, our intuitions about particular cases make up the basic data of ethical theory. However, most theories leave open the possibility that intuitions are mistaken, and that theory can provide the tools for evaluation of the intuitions themselves. The intuitionist approach does not appeal to an underlying theory.

(34) It should also be noted that some of these principles have also, historically, been grounded in the natural law tradition, which holds very roughly that moral principles have an authority based on human nature, or what's natural for human beings. This, of course, doesn't distinguish the natural law approach completely. While it needn't have a religious basis, it often does. Indeed, the classical presentation of the approach is in

Thomas Aquinas's *Summa Theologica*, where he holds that natural law is the "divine" law that is discoverable by human reason. However, more recent presentations of this approach can be found in John Finnis's work, particularly Finnis (1980). See also George (1993).

(35) This case is a modification of one presented in Kagan (1989).

(36) See, in particular, Scheffler (2000).

(37) To use Kagan's terminology, one might be a deontologist at the factoral level in terms of recognizing constraints, but consequentialist at the foundational level.

(38) Gilbert Harman, in discussing morality's conventionalism and the "intuitive" principle that harming is worse than not helping, writes that our intuitions have been shaped by social reality, that is, a kind of tacit understanding that the poor will not harm the wealthy in exchange for a weaker commitment to a norm of mutual benefit: 'The rich and powerful do not need much help and are often in the best position to give it; so, if a strong principle of mutual aid were adopted, they would gain little and lose a great deal…. On the other hand, the poor and weak might refuse to agree to a principle of non-interference or noninjury unless they also reached some agreement on mutual aid. We would therefore expect a compromise…. In the present case, the expected compromise would involve a strong principle of noninjury and a much weaker principle of mutual aid—which is just what we now have' (Harman 1977: 110–1).

(39) See also Tamara Horowitz's excellent (1998). Here Horowitz attacks Quinn's reliance on thought experiments alone to reveal moral norms, but the case she makes has broader implications. Her view is that prospect theory in psychology accounts for our intuitive distinction between doing and allowing, and that this distinction is, therefore, not morally significant. Prospect theory is developed by Daniel Kahneman and Amos Tversky (1979). Horowitz writes in her essay, 'I do not see why anyone would think the distinction is morally significant…If the distinction is not morally significant, then Quinn's thought experiments cannot play the role in his argument that he intends for them to play…. But to the extent that the intuitions elicited by Quinn's thought experiments are explained by prospect theory, they are not moral intuitions at all' (1998: 381). For responses to Horowitz, see Kamm (1998) and van Roojen (1999).

(40) For a response to Kagan, see Malm (1992).

(41) See also his essays in Gauthier (1990), particularly 'Justice as a Social Choice' and 'Bargaining and Justice'.

(42) Mintoff points out a variety of other problems for Gauthier's account and then offers his own contractarian account of rational choice; see Mintoff (1997).

(43) And, indeed, David Richards (1971) does this.

(44) See Sandel (1982).

(45) See Pateman (1988).

(46) See Kittay (1999).

(47) See e.g. Pettit (1993, 1999); McGinn (1999). For a defence of Scanlon against this type of criticism, see Ridge (2001).

(48) For more on this problem for Gauthier's contractarianism, see Morris (1991).

(49) For a more detailed discussion of this condition in Hursthouse's account, see Driver (forthcoming).

(50) I have adapted this case from some of the cases one sees in the literature on actualism, which is a view that holds that the agent ought to consider what he will actually do in determining what he ought to do. For more on this issue, see Jackson and Pargetter (1986).

(51) I make this pointin Driver (2003).

(52) On Spelman's view, too much focus on the care perspective can lead people to ignore the various ways in which women have mistreated each other, and the significance of that to doing feminist moral theory.

Julia Driver

Julia Driver is a Professor of Philosophy at Washington University in St. Louis. She received her PhD in Philosophy from The Johns Hopkins University. Her main areas of research are normative ethics, moral psychology, and metaethics. She has published articles in journals such as the Journal of Philosophy, Philosophy & Phenomenological Research, Nous, Philosophical Studies, Utilitas, Ethics, and the Australasian Journal of Philosophy. She has written three books: the latest, on Consequentialism, was published in 2012 in Routledge's New Problems in Philosophy series.