DEONTOLOGY

1 *Introduction*

What determines which actions are morally required? According to Act Consequentialism (AC),

the right action is the one that produces the most value (the best state of affairs, which may

include the act itself). Deontology denies this. One of our foci is to contrast deontology with

consequentialism, and clarify the debate between them. Thus, in addition to defending our

deontological view, we devote attention to consequentialist positions.

Railton (1988, 113) refers to AC as 'objective consequentialism.'[1] He contrasts it with

'subjective consequentialism' – '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. One is behaving as subjective consequentialism requires …

to the extent that one uses and follows a distinctively consequentialist mode of decision making,

consciously aiming at the overall good and conscientiously using the best available information

with the greatest possible rigor' (1988, 113). 'Simple Consequentialism' (SC) is our term for the

combination of subjective and objective consequentialisms. (In 3.1 we discuss other varieties.)

Deontology, in contrast to SC, claims that the production of good is not the only

fundamental morally relevant consideration: agents may be permitted, and even required, not to

maximize the good. There is much debate about details, but the basic distinguishing features of

deontology standardly fall under three rubrics.

2 *Basic Features of Deontology*

2.1 *Constraints*

Deontologists characteristically hold that we must not harm people in various ways. We should

not lie, kill innocent people, or torture anyone. These prohibitions *constrain* us in what we may

do, even in pursuit of good ends. Deontologists differ in how stringent these constraints are.

Some think them absolute. Roman Catholic moral theology has traditionally held that one may

never intentionally kill an innocent person. Kant infamously argued that it would be wrong to lie,

even to prevent murder. Other deontologists have held that, though constraints are always a significant consideration, they may be overridden, especially if that is the only way to avoid catastrophe. Either way, deontology sometimes *requires* agents not to maximize the good. Whilst, of course, any moral requirement restricts us in what we are permitted to do, we shall use the term 'constraints'[2]: to refer to moral restrictions that may require one not to maximize the good, where these restrictions do not stem from our special relationships to others. These latter restrictions fall under a separate category: duties of special relationship.

2.2   *Duties of Special Relationship*

Many of our duties stem from special commitments to others. Some commitments are explicitly undertaken, such as promising. Some are tacit – as in commitments to friends. Some are not voluntarily acquired – consider commitments to parents. Like constraints, the responsibilities that come with relationships curtail our freedom of action even when we could maximize the good by shirking them. John might benefit more from my help than will Jane, but if I have already promised Jane to help her, and I cannot help both, then it is Jane I ought to help. Duties of special relationship differ from constraints in that they are owed, by their very nature, only to those to whom we stand in such relationships, whereas there are constraints against torturing or unjustly killing *anyone*.

2.3   *Options*

Given the amount of suffering in the world and the disparities in wealth, to follow SC and maximize the good would require enormous sacrifice from anyone with more than a minimal standard of living. SC may thus seem too demanding. Many deontologists suggest that our duty to help others is limited. There is some point, though its location is hard to determine, at which agents have done all that duty demands. At that point they have an *option* to decline to do more[3]. We admire those who make the extra sacrifice, but it is supererogatory – more than morality requires. SC leaves no conceptual space for supererogation.[4]

Deontologists don't deny that morality can be demanding. We may be obliged to make significant sacrifices – even of our lives – rather than breach a serious constraint or betray a friend. And we have a duty to do good. But, unlike SC, most deontologists see this latter duty as limited.

2.4   *Agent-relative and Agent-neutral Theories*

Traditionally, SC and deontology are distinguished by their differing accounts of the relation between the right and the good. SC holds that the good determines the right – the amount of goodness produced by an action is the sole determinant of its rightness – whereas the deontologist denies this, holding that other considerations are relevant. More recent writers[5] distinguish between the two in terms of agent-relativity and agent-neutrality, claiming that SC is an agent-neutral theory whereas deontology incorporates agent-relative elements.

The distinction between the agent-neutral and the agent-relative may be introduced by reference to reasons for acting[6]. Roughly, someone's reason is agent-relative if, at base, there is reference within it to the agent. For example, egoists hold that each of us has reason to promote only *her own* welfare, whereas utilitarians believe each of us has reason to promote the general welfare. Note that both varieties of reason apply to us all, but agent-neutral reasons incorporate an added element of universality: to say that each of us has reason to promote the general welfare is to say that each of us has reason to pursue the *common* aim[7] of promoting the general welfare (and this requires that any person sacrifice her welfare if that will increase the general total), whereas according to egoism, each of us has a *distinct* aim: I have reason to pursue my welfare, you yours.

How does this distinction mesh with that between SC and deontology? SC holds that all moral reasons are agent-neutral, whereas deontology denies this. According to SC we each have reason to maximize the good, and morally speaking this is all we have reason to do. We have one common moral aim: that things go as well as possible. Someone may object that we

have distinct aims because my aim is that *I* maximize the good, and your aim is that *you* do so. Perhaps there are circumstances in which *my* maximizing the good does not result in the good being maximized. But this is to misread SC. Suppose I can directly produce ten units of good or five, and in the former case you will directly produce zero whereas in the latter you will produce six. SC prescribes that I directly produce the five, since the total produced will then be greater. SC cares not about who produces what directly, but about what is produced overall.

Deontology, by contrast, maintains that there are agent-relative moral reasons. Duties of special relationship are obviously agent-relative. That she is *your* daughter gives you special moral reason to further her interests. On this view, I am required to care for *my* family, you for *yours*: we have distinct aims. Contrast this with an SC view on which parental care-giving is valuable. On this view we have the common aim of promoting parental care-giving – which requires that I neglect my own children if I can thereby increase the total amount of parental care-giving.

Constraints are also agent-relative. Suppose I can only prevent you killing two innocents by killing one myself. Those deontologists who advocate an absolute constraint against killing the innocent forbid my killing the one (they also forbid, of course, your killing the two, but we are assuming here that you will ignore this proscription): I have overriding moral reason (a distinct moral aim) not to kill anyone *myself* (as you should aim not to kill anyone *yourself*). Thus although you will do wrong in killing the two, I should not kill the one in order to prevent you. SC, by contrast, holds that, ceteris paribus, I should kill the one: killing innocents is bad, so I have an agent-neutral moral reason to contribute to the common aim of minimizing the killing of innocents.

Options need not be agent-relative in their formulation. They simply permit us not to maximize the good. But their standard rationale is agent-relative. Each of us is morally permitted to give special weight to *our own* interests.

There seem to be two ways of distinguishing between agent-relative and agent-neutral moral theories. On the one hand, theories prescribe aims, and these can be common or distinct. By this criterion, a moral theory is agent-neutral exactly if it prescribes common aims, and is agent-relative otherwise. On the other hand, a theory is agent-neutral just in case it countenances only agent-neutral moral reasons, and is agent-relative otherwise. SC is agent-neutral, and deontology agent-relative, on either account.

Common-sense morality[8] (CSM) acknowledges special obligations, constraints and options. Thus deontology is closer to CSM than SC in this regard. Those advocates of SC who are radical reformers[9] claim that CSM is mistaken here. But many moral theorists hold that we cannot ignore our common-sense moral intuitions, seeing them as a key source of evidence. Other non-deontological theories, then, including other forms of consequentialism, endeavor to achieve a closer fit with our moral intuitions by allowing room for agent-relative considerations. We turn next to discussing some of these theories.

3   *Non-Deontological Theories*

3.1   *Consequentialisms*

According to AC, the right action is the one that maximizes good ('right equals best'). SC supplements this with a decision procedure that has us 'consciously aiming at the overall good' (Railton 1988, 113). SC, then, is apparently direct in the sense that one should employ the criterion of right action in deciding what one ought to do.

But there is a complication. Regan (1980, 264-65, note one) offers the following example. You must choose between acts *f* and *g,* where *f* has an even chance of producing zero or ten (objective) utiles, and *g* is sure to produce 9 utiles. Unbeknownst to you, *f* will produce 10 utiles, and is thus best, and hence the right act by AC's criterion. But surely you ought to *g.* We shall interpret 'consciously aim at the overall good' as 'consciously aim at maximizing expected objective utility'. We think of 'ought' as action-guiding; thus when we speak

of what the agent ought to do, we are referring to the output of the recommended decision procedure when correctly followed. Thus what the agent ought to do according to SC is what, on the basis of the information available to her, she calculates will maximize expected objective utility, where no calculation error is made but where her information may be less than full. We shall leave in place the AC criterion of rightness. Thus in Regan's case, according to SC, what you ought to do is the wrong thing. You cannot, however, know in advance that it is wrong. SC, then, maintains that 'right equals best', even where you cannot know what is best. What you ought to do is epistemically accessible, thus what you ought to do may be wrong. But what you ought to do is never something that you know at the time to be wrong.

Some authors[10] refer to a consequentialist theory as 'direct' just in case it is a form of AC. Rule consequentialism (RC) is an example of a consequentialist theory that is not direct in this sense. According to RC: 'An act is wrong if and only if it is forbidden by the code of rules whose internalization by the overwhelming majority … has maximum expected value' (Hooker 2000, 32). RC assesses rules, but not acts, in terms of their contribution to the good. Wrong actions are those that violate the rules. Thus an act may not be wrong and yet fail to produce the best ('right does not equal best') since the optimal rules must be, for example, simple enough to learn and sufficiently appealing that people will generally follow them. So they will often lead us to do less good than we could. Rules that would be fine for angels might be disastrous for humans. And even if humans could be trained to follow them, the cost of inculcating them might be too high.

AC theories, whilst direct in the foregoing sense, may be psychologically indirect: they may tell you not always to think about the (expected) good in deciding what to do (i.e., they may not be subjective consequentialisms in Railton's sense), because you may produce less good if you are obsessively concerned with its production.

Railton is an AC who advocates such psychological indirection. His 'sophisticated

consequentialist is someone who has a standing commitment to leading an objectively consequentialist [i.e., AC] life, but who need not set special stock in any particular form of decision making and therefore does not necessarily seek to lead a subjectively consequentialist life' (1998, 114). Indeed, it may be that a sophisticated consequentialist (Sophisticated C) 'should have (should develop, encourage, and so on) a character such that he sometimes knowingly and deliberately acts contrary to his objectively consequentialist duty' (1998, 121) – i.e., unlike the SC, the Sophisticated C *can* know in a particular case that what he ought to do is the wrong thing.[11]

Consequentialist theories may be indirect in both senses. Consider RC. Given that the optimal code has been internalized as part of a 'shared conscience' (Hooker 2000, 2) the agent should, it seems, generally follow her conscience rather than worry about the rules. In the case of RC, we interpret the question of psychological indirection as asking whether we ought to worry about right and wrong when deciding what to do[12], hence RC is psychologically indirect: the agent should not always consciously employ the criterion of right action in deciding what to do.

Although RC and Sophisticated C are both psychologically indirect, they may differ on how to think in morally tricky cases. According to Sophisticated C, when faced with a morally tricky decision where deliberation is in order, you should often *not* focus upon the right (but upon, say, your spouse). But for RC, in such cases, it is plausible to maintain that you *should* focus upon the right – either by wondering what your current rules tell you to do in this case, or by wondering whether your current rules are the best set.

Motive consequentialism (MC), as we shall understand it[13], is similar to RC vis-à-vis indirection.[14] The right act need not be best, but is in conformity with the best set of motives[15]. And, assuming one has internalized this set, one should in general simply follow it without worrying about rightness. MC and Sophisticated C differ in their criteria of right action, although

both claim that virtuous agents act in accord with the best motives.

Our classificatory efforts are summarized in table I.

Table I

	Act is wrong if not best (AC theories)	Sub-optimal act need not be wrong
Psychologically direct	SC	
Psychologically indirect	Sophisticated C	MC, RC

In (2.4) we characterized agent-relative moral theories in terms both of reasons and aims. We noted that deontology is agent-relative on both accounts, and SC agent-neutral. How are MC, RC and Sophisticated C to be classified? If we think of aims as the outputs of decision procedures, all three are agent-relative by the aim criterion.

RC incorporates, for example, a fairly simple rule against killing the innocent, because the adoption of a more complicated rule that allowed killing in pursuit of the good would be harder to follow and would undermine our valuable reluctance to kill. (Such prohibitions will not be absolute. Agents are permitted to breach them when catastrophe threatens.) There will also be rules that require us to devote time and energy to looking after friends and family.[16] These rules give each agent a distinct aim. I have the aim that *I* not kill the innocent, and look after *my* family etc.; you have the aim that *you* not kill the innocent, and look after *your* family. And these rules do not permit their own violation in order to promote greater conformity to them.

Similar remarks apply, *mutatis mutandum*, to MC and Sophisticated C.

Turning to the reason criterion, we assume that, for MC and RC, one has most reason to do the right thing. Thus MC and RC are agent-relative by this criterion also: a rule (motive) against killing the innocent is part of the best set (where, recall, this is determinative of rightness), thus I have strong moral reason not to kill the innocent *myself* (even to minimize the

number of such actions).

The case of Sophisticated C is not so straightforward. Does one have most reason to do the right thing or to follow the best motives? If the former, Sophisticated C is agent-neutral by the reason criterion (one has moral reason only to maximize the good); if the latter, agent-relative (the Sophisticated C has the same best motives as the MC).

If, then, these theories are genuinely consequentialist, it might appear that the agent-relative/agent-neutral distinction is not the apt way to draw the distinction between deontology and consequentialism. On the traditional distinction, by which consequentialism, but not deontology, claims that the right is determined solely by the good, at least AC (and hence Sophisticated C), although not RC or MC, is consequentialist.

There is another sense, however, in which all three theories are agent-neutral. We can assess an action's value from some particular person's perspective – we can ask, for example, whether it is bad for her. But we can also assess its value impersonally. For instance, pain is bad, regardless of whose it is. Claims about impersonal value make no fundamental reference to any particular agent, and so in this sense impersonal value is agent-neutral. Each form of consequentialism assesses something, at its base, in terms of impersonal value. But what they assess varies. AC assesses acts; MC assesses motives; Sophisticated C assesses both acts and motives; and RC assesses rules. As Hooker notes: 'the agent-relativity in RC is derivative. Agent-relative rules are justified by their role in promoting agent-neutral value' (Hooker 2000, 110). Similarly, the agent-relativity in Sophisticated C and MC is derived from considerations of agent-neutral value. Deontology, by contrast, holds that some agent-relative considerations are *underivatively* relevant. They have weight in their own right, not merely in virtue of their serving some further purpose.[17]

3.2   *Non-Consequentialist Non-Deontological Theories*

Some non-consequentialist theories have the same structure as RC, but offer non-consequentialist criteria for selecting the rules. Copp's society-centered theory, for example, sees a rule as 'justified in relation to a society just in case the rule is included in the moral code that the society would be rationally required to select, in preference to any other code, to serve as its social moral code' (Copp 1997, 190). It is practically rational for individuals or societies to select, roughly, what would best satisfy their needs and further their values. Such a theory will generate rules rather similar to those endorsed by RC. However, since the test for what rules are acceptable does not appeal to agent-neutral value, the theory is not consequentialist. But that does not make it deontological, since the moral force of any agent-relative considerations is only derivative.[18]

Unlike Copp's theory, Scheffler's (1994) theory makes room for underivative agent-relativity. It agrees with SC's rejection of constraints but incorporates options, and justifies these by appeal to the cost to *the agent* of maximizing the good. In refusing to give the personal perspective any *moral* weight, consequentialism does not reflect the *natural* weight that agents give to *their own* projects, friends, family etc. Scheffler's theory is thus not consequentialist: it allows that agent-relative considerations have fundamental moral weight in justifying options. Yet arguably it is not a deontology. The latter, strikingly, sometimes *requires* us not to maximize the good, but Scheffler merely *permits* this, when it would significantly damage our concerns.

Having classified various non-deontological theories, we turn now to classifying deontologies.

4   *Rossian Deontology*

Ross (1930, ch. 2) claims that there are several distinct underivative agent-relative moral considerations, which he formulates as a list of basic principles or duties. These include agent-relative duties of promise-keeping, gratitude, reparation, and not harming others[19]. In addition, he agrees with SC that there is an agent-neutral requirement to promote the good (which

includes, for Ross, justice). These duties are only *prima facie* (or, as we prefer, *pro tanto*) since, though each is relevant to determining what is right, they can conflict. If keeping a promise will harm someone, for example, to determine what is right, the duty to keep the promise must be weighed against the duty not to harm, where this weighing is governed by no higher rule – it requires discernment and judgment.

Whilst other deontologists, such as Kant and Scanlon (see below), agree with Ross that there are a number of basic principles, they see them as basic only in the sense that they are not instances of some more general principle. For instance, all three might agree that the duty to pay one's debts is not basic because it is an instance of the duty to keep promises, whereas the latter is basic since it is not itself an instance of some more general duty. For Kant and Scanlon, however, but not Ross, even such basic principles rest on a common foundation (although Kant and Scanlon disagree as to what this is): there is a test that principles must pass – a test, furthermore, that is claimed to be in some sense definitive of morality. Ross denies that there is such a test. (SC agrees with Ross in this: there is no test that SC's fundamental principle, maximize the good, must pass.)

Ross also claims that the *prima facie* duties are self-evident. By this he means (roughly) that they stand in no need of justification[20], and we can see their truth directly, without reasoning from further premisses.

In formulating his principles, Ross assumes that if a consideration is fundamentally morally relevant in one case it is relevant in the same way in all cases. If we have a fundamental *prima facie* duty not to harm, then the fact that an act will cause harm is invariably a moral reason not to do it, though not necessarily an overriding one. Harmfulness has invariably negative moral valence. Apparently Ross reasons thus. Any feature of an action may be morally relevant to its rightness, but many features are merely derivatively relevant. That it is Tuesday is morally relevant if I have promised to do something on Tuesday, but its relevance derives from

the content of my promise. What is fundamentally relevant, however, cannot derive its moral force from elsewhere, and so must have it essentially. Its valence will not vary.

This argument is, however, invalid. A moral consideration may be basic, in that wherever it counts its moral force is underivative. Yet its force may be conditional on the presence of other features – it might not count in all cases. Take promise-keeping. Ross claims that my having promised to do an act always counts in favor of doing it. But this is mistaken. Promises extracted by fraud or force are null and void, as are promises to do something immoral. Suppose I promise to perform a contract killing. It is implausible to hold that, though I ought not to do it, all things considered, the fact that I promised gives me *some* moral reason to do it. The duty to keep promises is not derivative – when we have reason to keep a promise, there is no more basic moral reason that explains why – but it is conditional.[21].

It may seem, however, that Ross can address this concern. We could arrive at a consideration that has invariant valence by simply adding the relevant conditions. Particularists (who hate a principle), however, disagree.

5   *Particularism*

To what extent is morality codifiable? Some hold that moral theory should refine and qualify our moral principles so that a verdict can always be 'read off' from them. Ross and the particularist agree that there are no such verdictive principles. What of non-verdictive principles? Are there, for example, non-normative features with invariant valence? The particularist says not. For example, it might be claimed that there is always a moral reason not to lie. But Dancy (1993a, 60-61) raises the case of children's games in which lying is part of the fun. Lying in these contexts, he claims, does not carry negative weight. The particularist's general claim here is that any non-normative feature varies in valence according to context.

One response to this general claim is to increase the complexity of the non-normative features. Perhaps lying has universally negative valence except in contexts in which all relevant

parties tacitly acknowledge its acceptability. One worry here, however, is that acceptability is itself a normative notion. Second, there may be other exceptions to the principle that lying counts negatively – indeed, there may be no finite list of exceptions that suffices.

But Ross, we think, largely accepts this line, at least tacitly: apart from the case of (perhaps) promise-keeping, his principles claim invariant valence only for normative notions such as justice, gratitude, loyalty and reparation. Ross's principles, then, are usually couched in normative terms, and thus it requires moral sensitivity to determine whether an act falls under them.

Principlists might retort, however, that (1) there must be non-normative features with invariant valence because we cannot make moral judgments without appeal to them; and (2) consistency is essential to morality, and to be consistent is to follow principles or rules.[22]

The response to (1) is to note that we test whether, say, lying under certain conditions has universally negative valence by searching for counter-instances. But the very possibility of such a search shows that we can tell whether lying is relevant in a particular case without appeal to our principle about lying. When we come across circumstances, real or imagined, that force us to qualify a principle, we recognize that the qualification is required. And this recognition, on pain of regress, is not achieved by appeal to some further principle. Furthermore, even if there are principles with finite numbers of exception clauses, the particularist claims that we can never know that we have listed them all. Principles are epistemically redundant.

The response to (2) is to recall one of the Wittgensteinian rule-following considerations. Suppose the principle is 'lying has universally negative valence'; then, in order to apply it, one must be able to determine whether a novel case is a case of lying. But to do so consistently, according to the view that consistency requires appeal to principles or rules, requires appeal to some further rule. Even if such can be formulated, its consistent application will require appeal

to further rules, and so on. Eventually, there must be brute application without appeal to rules. And this will vitiate claims to consistency (on this view of consistency). Thus either we are inevitably inconsistent, or consistency is not a matter of rule-following.

According to the particularist, then, reasons function holistically: no consideration is uninfluenced by its surroundings – the relevance of any feature may vary according to context. Do we, as Rossians, agree? Not if the particularist maintains that *all normative* features have variable valence. If the notion of a promise is a normative notion, then we might agree that promise-keeping is a normative feature that can vary from having positive valence to being irrelevant (though we doubt that there are cases where the fact that I made a promise counts *against* keeping it); but we find it implausible that there are cases in which features such as justice or loyalty are morally irrelevant or even negative. Justice and loyalty are thick moral concepts. These are, roughly, those associated with the virtues and vices. And these, we contend, have invariant valence.[23] The particularist view that anything may count (or not), and in ways that cannot be specified in advance, seems to have no way of accounting for the moral centrality of the thick moral concepts beyond noting that they are more frequently relevant than others.

But have we not conceded too much to the particularist? It might appear that not only are there non-normative considerations that have invariant valence, but that there are actually such considerations that always make an act, say, wrong. We agree, for example, that gratuitous torture is always wrong. But we claim that the notion of gratuitous torture cannot be spelt out non-normatively. One might try: inflicting pain on another for no reason. But the difficulty is that reason is a normative notion. So how about: inflicting pain on others solely for one's own pleasure is always wrong. Again, we agree. But the proscription is tantamount to: inflicting pain on others where the only reason for doing so is one's own pleasure. And, again, this adverts to the normative notion of a reason.

Some form of Rossian deontology seems to us to strike the correct balance between principlism and particularism. Having defended Ross from the particularist, we now turn to alternative foundations for deontology and the attack from principlism.

6   *Alternative Foundations for Deontology*

Rossian deontology seems to have a number of drawbacks. First, there is a diversity challenge: in uncovering the fundamental moral principles, Ross appeals solely to our reflective convictions, so what do we say to those whose reflective moral beliefs differ significantly from ours? Second, we have a reasons hurdle: how do we argue with those who doubt that moral considerations are reasons? Third, there is the no algorithm difficulty. There are disagreements about what is right, both because people cannot agree about the weight to be given to competing considerations, and because there are disagreements about how to apply a principle. We may agree that harming others is *pro tanto* wrong, but disagree both about how this weighs against other considerations in a particular case, and about what constitutes harm. Ross offers no algorithms for deciding difficult cases. Finally, rather than a unified account of the nature of obligation, Ross offers an irreducible list of disparate fundamental considerations – the unity problem.

Other deontological theories might appear to do better in addressing these difficulties. Some, such as Scanlon's (see 6.2), arrive at the content of morality by considering what principles people have reason to agree to. Others, such as Kantianism, ask what principles could be universally accepted.

6.1   *Kant*[24]

Kant sees morality as a species of practical rationality, and offers a test of the latter: the categorical imperative (CI) test. Actions that fail this test are, he claims, wrong. Crucial to the test is the notion of a maxim. We act with certain aims (which we might not have consciously formulated), and these can be specific or general. Maxims are general aims. Thus my maxim

may be: make lying promises (i.e., ones I intend not to keep) whenever it benefits me. The CI

test asks first on what maxim I propose to act, and then enquires whether this maxim is one that

I could will to be a universal law. Here is a rough illustration[25]: the maxim to make lying

promises whenever it benefits me cannot be universally willed because its universal adoption

would lead to the demise of the very practice on which it relies – namely the practice of

promising. Hence making lying promises for my own benefit is wrong. (One issue here is

whether the fact that the maxim on which you acted cannot be universally willed is even relevant

to the issue of why the action is wrong.)

How exactly the CI test is to be understood and what it would rule out are matters of

scholarly dispute[26]. But there is general agreement that Kant's ethics has a deontological

structure. The test yields constraints, for agents are forbidden, on an alternative formulation of

the test, to treat others *merely* as a means. Exactly what this entails is again in dispute, but it is

intended to rule out such things as lying and killing the innocent even to minimize lying and the

killing of innocents by others. To kill an innocent yourself to prevent other killings, for example,

would be to use your victim as a means to minimize victimization. From SC's perspective, these

constraints forbid one to maximize the good. Kant's system also admits options: we have only a

limited duty to help others[27].

The rationale for Kant's test lies in a certain conception of rationality. If something is a

reason for one agent then it must be capable of being a reason for all. Thus a maxim is not a

good reason for action unless it is one on which all agents can act. Any maxim that could not

consistently be followed by all, or could not consistently be willed as one that all should follow, is

not rationally acceptable – it fails to show respect for the autonomy of all other rational agents.

Kant's theory seems to overcome the supposed drawbacks of Rossian deontology. It

meets the diversity challenge because Kant's test for right action is a purely formal one,

appealing only to what can be willed consistently. It does not presuppose any substantive

evaluative or deontic claims. It leaps the reason hurdle by claiming that it would be irrational to act on a maxim that could not be universally adopted. It avoids the unity problem because the test offers a unified underpinning to our disparate duties. Finally, it goes some way to surmount the algorithm difficulty, in claiming that certain kinds of consideration are morally decisive. Some duties – the duty not to lie[28], or kill the innocent, for example – are held to be absolute; i.e. they can never be overridden by other moral considerations. Clear and unequivocal moral guidance is, however, here bought at a high price. For the claim that it is always wrong to lie, even to save a life, runs counter to most people's moral intuitions.

Equally counter-intuitive is Kant's claim that only other persons have moral claims on us – non-rational creatures have no independent moral standing (for instance, the fact that an action would cause suffering to an animal is itself no reason not to do it, according to Kant). These and other well-known objections to Kant's theory prevent it fulfilling its ambitious programmed.

6.2   *Scanlon*

Whereas Kant asks of a principle whether rational agents could universally will it, Scanlon asks whether reasonable persons could reject it. (Scanlon sees his position as continuing the social contract tradition, hence the name 'contractualism'.) On Scanlon's view:

> An act is wrong if its performance under the circumstances would be disallowed by any set of principles for the general regulation of behavior that no one could reasonably reject as a basis for informed, unforced general agreement. (1998, 153)

Our aim is 'to find principles [for the general regulation of behavior] that others who share this aim also could not reasonably reject' (Scanlon 2002, 519), where someone may only object to some proposed principle if its general acceptance would place excessive or arbitrary burdens on her. Whether the objection constitutes grounds for reasonable rejection depends on whether there is a comparable principle available that is not subject to similar objection (1998, 205).

Scanlon distinguishes between narrow and broad morality (1998, 6-7, 171 ff). Narrow morality is his central focus: it is this that concerns wrongness as defined above, and is captured by the phrase 'what we owe to each other' (1998, 7). While Scanlon agrees, for instance, that 'pain – whether that of rational creatures or nonrational ones – is something we have prima facie reason to prevent, and stronger reason not to cause' (1998, 181), there is (ceteris paribus) more reason to respond to the pain of a rational creature: not only is the rational creature's pain bad, but in addition 'we may owe it to him to help relieve it.' The fact that A-ing is wrong is a reason not to A that augments the other reasons against A-ing (1998, 11) (where wrong-doing cannot be committed against nonrational creatures).

Scanlon is sympathetic to deontology[29], so how does his view differ from SC? SC requires us to aggregate value across persons, which requires significant sacrifice on the part of a few in order to produce a relatively small benefit for each of the many if more good is produced thereby. Contractualism, by contrast, holds that 'the justifiability of a moral principle depends only on various *individuals'* reasons for objecting to that principle and alternatives to it' (1998, p. 229; italics his). And this, Scanlon thinks, is sufficient to block objectionable aggregation.

Aggregation is counter-intuitive if values are not on a par, and contractualism captures this. Suppose we can rescue an electrician in a TV transmitter station, but only by switching off the transmitter and depriving millions of World Cup football. No viewer could reasonably reject the principle that one must 'save a person from serious pain and injury at the cost of inconveniencing others or interfering with their amusement … no matter how numerous these others may be' (1998, 235). Each viewer's complaint is so trivial that, no matter how many are affected, the electrician should be rescued.

A consequentialist who sees values as lexically ordered, however, would aggregate only when values are on a par. But Scanlon denies any appeal to aggregation. Suppose there are

two groups of people, the second more numerous than the first. Suppose, further, that I am

morally required to save at least one group but cannot save both. On aggregative grounds I

should obviously save the second group. However, since all the individuals apparently have the

same complaint, none of them, it seems, can reasonably reject a principle that permits the

saving of either group. Scanlon rightly sees this as counter-intuitive. But he cannot appeal

directly to aggregation on pain of turning consequentialist and rejecting the central importance

of individual complaints:

> It therefore seems that as long as it confines itself to reasons for rejection arising from
>
> individual standpoints contractualism will be unable to explain how the number of people
>
> affected by an action can ever make a moral difference (1998, 230).

Scanlon saves contractualism here by noting that a person from the second group can protest

against a principle permitting the saving of either group that, were she not present, it would still

be permissible to save either group. Thus her presence apparently makes no difference – it is

as if her life has no 'moral significance' (1998, 232). But her life, she protests, has the same

moral significance as everyone else's. And this is a complaint from an individual standpoint.

Does this appeal to individual standpoints make Scanlon's system deontological?

Scanlon intends that the individual complaints for rejecting principles be agent-relative (for

instance, a complaint that the adoption of a principle would not maximize the good is not an

individual complaint), thus the theory acknowledges underivative agent-relative considerations.

On the other hand, Scanlon's theory has agent-neutral elements. Given several competing

principles, Scanlon's theory would presumably require us to rank-order the complaints against

them in accord with their seriousness, and then select the principle that suffers the fewest

complaints at the most serious rank: we have the common aim of minimizing the number of

complaints at the most serious rank. We can also ask, of course, whether the principles that are

so selected will be agent-relative or agent-neutral.

For example, each of us has reason to want a principle concerning the wrongness of taking of human life.[30] But what form should this principle take? We might begin with 'killing people is wrong'. But, as Scanlon notes, 'what about self-defense, suicide, and certain acts of killing by police officers and by soldiers in wartime? And is euthanasia always strictly forbidden?' (1998, 199) He continues:

> The parts of this principle that are clearest are better put in terms of reasons: the fact that a course of action can be foreseen to lead to someone's death is normally a conclusive reason against it; the fact that someone's death would be to my personal advantage is no justification for aiming at it; but one may use deadly force when this seems the only defense against a person who threatens one's life; and so on.

Perhaps Scanlon has in mind a principle along the following lines:

> *Principle K:* If A sees that X can be foreseen to lead to someone's death, then, in the absence of special justification (such as self-defense), A must not do X.

Whether K is reasonably rejectable depends, of course, on what counts as 'special justification'. This is a phrase lifted from Scanlon's principle F (1998, 304), and a key question is whether the fact that my killing one would save several others from being killed by another is such a justification. If so, then K is not a deontological constraint, but is, rather, consistent with an agent-neutral prescription to minimize killing – ordinarily the best way to do this is not to kill anyone yourself, but there are exceptional circumstances where this is not so.

There are parallels here with the case of saving the more numerous of two groups. Suppose by killing Jane I can thwart your effort to kill John and Joe. And suppose K is interpreted as forbidding this. Then John (or Joe) can complain that it is as if his life has no 'moral significance'. There is some question, then, as to whether Scanlon's theory incorporates a deontological constraint against killing. And to the extent this 'moral significance' argument can be generalized, there may be similar questions raised about other principles. Additionally,

Scanlon departs from traditional deontology in not seeing special obligations to one's friends and family as *moral* obligations (1998, 162).

Suppose, however, that Scanlon's contractualism is a deontology: how does it compare with Rossian deontology? The Rossian agrees that:

> Principles … are general conclusions about the status of various kinds of reasons for action. So understood, principles may rule out some actions by ruling out the reasons on which they would be based, but they also leave wide room for interpretation and judgment (1998, 199).

So, for example, the Rossian concurs that killing solely for personal gain is wrong. And Scanlon manifests particularist tendencies when he notes (1998, 51) that some feature may be a reason in one context, but not in another.

Furthermore, Scanlon appeals to Rossian reasons in his discussion of broad morality: as we have seen, he thinks that 'pain – whether that of rational creatures or nonrational ones – is something we have prima facie reason to prevent, and stronger reason not to cause' (1998, 181). But what about narrow morality? In Scanlon's view: 'What is basic to contractualism … is the idea of justifiability to each person (on grounds that he or she could not reasonably reject)' (1998, 390, note 8). But what does the claim that justifiability on reasonable grounds is central add to the claim that reasons are central? Any rejection that is reasonable must be supported by reasons. Hence the digression through reasonable rejectability appears unnecessary[31]. Scanlon maintains that: 'What *makes* an act wrong are the properties that would make any principle that allow it one that it would be reasonable to reject ([such as] the needless suffering and death of [a] baby)' (1998, 391, note 21). But why can't we appeal, in Rossian fashion, to the reasons directly? A-ing would be wrong because it would result in the needless suffering and death of a baby, as opposed to: A-ing would be wrong because it would be in violation of a principle that it would not be reasonable to reject, and it would not be reasonable to reject this

principle because it would permit the needless suffering and death of babies.

Scanlon and the Rossian differ on their views concerning whether wrongness is itself a reason. The Rossian identifies wrongness with the presence of decisive negative moral reasons, whereas Scanlon thinks that wrongness is itself a decisive negative reason: 'The fact that an act is wrong seems itself to provide us with a reason not to do it, rather than merely indicating the presence of other reasons (although it may do that as well)' (1998, 11). Indeed: 'the fact that an action would be wrong constitutes sufficient reason not to do it (almost?) no matter what other considerations there might be in its favor' (1998, 148). We do not see this difference as redounding to Scanlon's advantage, however.

One worry is that Scanlon's account might lead to 'double-counting'. Suppose it would be wrong for A to kill B. Then the wrongness is a reason against the killing. But Scanlon acknowledges that there will be other reasons against the killing (such as B's reasonable complaint that it would unfairly harm him). The danger is that these other reasons against the killing will also be part of the reason why the killing is wrong and thereby get counted twice.

In his favor, perhaps Scanlon makes progress on the unity problem – at least when it comes to narrow morality.[32] Moral agents have the unifying aim of seeking principles of a certain kind. And narrow morality concerns 'what we owe each other', which is cashed out in terms of reasonable rejectability. But as we have seen, reasonable rejectability appeals to a wide range of reasons: Scanlon seems to allow an irreducible list of disparate fundamental considerations. We doubt, then, that Scanlon's claim to unity is any stronger than the Rossian's, particularly in light of the fact that the Rossian is considering broad morality.

On the issue of a decision algorithm for testing or generating moral verdicts, Scanlon readily allows that his 'principles … leave wide room for interpretation and judgment' (1998, 199).

Concerning the reasons hurdle, Scanlon takes it as given that 'everyone has reason to

seek and be guided by … principles [for the general regulation of behavior that no-one who shares the aim of finding such principles could reasonably reject]' (2002, 519) But this is not going to be accepted by those who doubt that moral considerations are reasons. Like the Rossian, Scanlon does not present arguments against such a skeptic.

Such objections only appear worrying, however, if we judge moral theories by more stringent standards than we deploy elsewhere. First, are there decision procedures in all other areas of knowledge? Second, the Rossian contends that there are a variety of moral considerations, and faces the apparent challenge of explaining what makes them all moral – but is there some underlying feature that explains what makes, say, all logical considerations logical? Third, the diversity challenge is no more of a problem in ethics than it is in many areas of enquiry. Finally, all normative enquiries face the reasons hurdle, including theoretical reason and logic. We can explore the status of any consideration's claim to be a reason. Why think the status of moral considerations is especially dubious?[33]

7   *Defending Deontology*

We turn now to the defense of our Rossian deontology against its two main contenders: consequentialism and virtue ethics. The latter we address briefly in section 7.4. There are many well-known objections to the former. SC, it is claimed, would over-burden us with calculations, and would demand too many sacrifices of us. MC and Sophisticated C may also seem unreasonably demanding in this latter sense: given the dispositions that others actually have, the best disposition for you to have might be to make continuous significant sacrifice on behalf of those in poverty. RC gets around this by asking, not what rules I should follow in the current situation, but what rules would be best if (almost) everyone accepted them – in which case (provided that those who accept rules tend to follow them), quite a modest level of self-sacrifice would eliminate avoidable suffering.[34] But RC suffers the charge of irrational rule-worship:[35] if the rules rest on considerations of value, how can it be insisted that it is wrong to override the

rules in pursuit of value?[36] And the list of objections continues. In sections 7.1 and 7.2 we focus

on social relations and autonomy, respectively, and maintain that deontology does better than

consequentialism with regard to them. In section 7.3, however, we are concessive to

consequentialism in our discussion of constraints, but we maintain that our view remains,

nevertheless, distinctively deontological.

7.1   *Special Relationships*

Deontology holds that there are underivative agent-relative moral ties between those who stand

in certain social relationships to each other. *Agent-relative*, because reference to the fact that *I*

am in the relationship is an ineliminable part of the reason why I should do something for the

other person: 'I owe it to her because she is *my* colleague, child, fellow-citizen etc.'

Underivative, because that reason does not rest on considerations about the general value of

people being in such relationships, or behaving in certain ways when they are. You have a right,

for example, to expect that I will give you a ride because I promised you, and not because of the

general utility of supporting the useful institution of promise-keeping. A *moral* tie, because it is

pro tanto wrong to be in breach of your special obligations to others.

On this matter, common sense concurs with deontology. Consider loyalty between

friends. It is not just that friends spend time with each other, support each other, and so on. In

addition, a friend has the right to expect *your* loyalty and support because she is *your* friend. If

you betray her, she has a moral complaint against you that no-one else has. Moreover, the

(tacit) acknowledgement of a moral tie between friends appears essential to friendship (as

placing oneself under an obligation is essential to 'successful' promising – i.e., promising where

none of the countervailing conditions are in play). Friends come through for one another;

someone who neither came through for you, nor believed she should, would not be loyal and so

would not be a friend.

If this is right, then consequentialism has a serious strike against it. Loyalty is essential

to friendship. Loyalty involves the recognition of an underivative agent-relative obligation to my friends. Consequentialism has no place for underivative agent-relative obligations; thus it has no room for friendship. But friendship, as is generally acknowledged by consequentialists, is an important intrinsic good. Consequentialism holds that the good is to be promoted; but here is a good that it cannot accommodate.

We have posed this as a problem for consequentialism generally, because although consequentialists of different stripes can respond differentially to this objection, they all deny the existence of agent-relative, underivative, moral obligations. SC simply denies that there are agent-relative obligations. RC, MC, and Sophisticated C deny that they are underivative. And self-effacing theories, which we introduce below, share SC's denial that there really are such obligations, while maintaining that it would be better if people believed there were. Let us look in more detail.

SC has no room for moral ties, and hence for friendship, because it has no place for agent-relative moral reasons. But can it accommodate a different account of friendship based on the idea that there are special psychological (and non-moral) bonds of affection between friends? We contend not. Even if we abandon the thought that we are *required* to favor friends, surely we must be *permitted* to favor them if our bonds to them are to be special. That is, we must be permitted to favor our friends even when we could do more good overall by not doing so. But SC denies us this permission: an act is wrong if it fails to maximize the good. All bonds are of equal importance: your bonds to your friends are of no more importance than bonds between others and their friends. You might have reasons of efficiency to tend to those nearest to you: it does more good for less effort to give flowers to *your* spouse. But reasons of efficiency do not allow room for *special* bonds of affection. Your spouse would not be heartened to discover that you, being a conscientious SC, have given him flowers only because love relationships are good, and this is the most efficient way to promote such relationships.

Does Sophisticated C do better? Assuming that a disposition to be loyal to those for whom one has special affection forms part of the motivational set that produces the best results, Sophisticated C requires loyalty in the sense that following its recommended decision procedure will result in loyal behavior. Thus virtuous agents are disposed to act loyally – even in some circumstances when disloyalty would produce more good. By Sophisticated C's lights, however, loyal action in such circumstances is wrong (there can be occasions when the virtuous agent knowingly does the wrong thing). Hence, like SC (both being forms of AC), Sophisticated C not only rules out moral obligations to friends, but acts of friendship are morally permitted only when they maximize the good. Sophisticated C might leave room for friendship, but only at the expense of endorsing immoral action.

How does MC fare? By MC's lights, one is permitted not to maximize value if that failure is in accord with the best motivations. Thus if the best motivational set contains friendly dispositions, it is permissible to favor friends. But if the MC asks herself why the fact that someone is her friend has moral significance, she will find herself ultimately appealing to considerations of the general good: she has a disposition to be loyal to Mary because the disposition to be loyal to friends is a good general disposition to have, not because of her particular relation to Mary – this relationship has no special moral importance for her.

Like MC, RC acknowledges that non-optimific acts need not be wrong, but like MC it fails to capture friendship because it maintains that preferential treatment of friends can be justified only by appeal to the general good:

> Moral requirements of loyalty are … needed … when affection isn't up to the job. … [S]pecial moral obligations towards family and friends can then be justified on the ground that internalization of these obligations gives people some assurance that some others will consistently take a special interest in them. Such assurance answers a powerful psychological need (Hooker, 2000, 141).

This does not yield genuine loyalty. Friends have moral reason not to let us down, and assurance is engendered in part by a belief that they will respond to this reason. (This is not to say that the only reasons here are moral.) But on RC, the *moral* reason for John not to let Mary down is the assurance that results from the internalization of a rule requiring the special treatment of 'friends', not anything special about his relationship with Mary.

RC's position is: given human psychology, it is best if each of us has special others who can be relied upon to reciprocate, thus the best set of rules takes account of this. And given the human tendency to feel special psychological connections to certain others, the least costly option is to inculcate a rule requiring their preferential treatment. However, if we now contemplate our reasons to favor them, we see that these reasons rest not on our putative special relationships with those others, but on the impersonal calculus of costs and needs.[37]

Some of our objections thus far have hinged on the possibility of an agent reflecting on her reasons for being loyal to friends and finding that these ultimately rest on considerations of the general good. Psychological indirection is of some help in addressing this concern: the virtuous agent does not consult the relevant theory each time she acts, be it towards friends or otherwise. But, we contend, were an agent to ask herself why she should be loyal to her friends, she would have to abandon her psychological indirection and would then see that her relationships with her friends are of no special moral importance to her.

At this point a further move is possible: the theory's psychological indirection could be strengthened so that the theory directs us not to believe it, thus placing the considerations we are directed to ignore permanently and completely out of reach. This would be to make the theory self-effacing[38]. Thus even though you should behave loyally to friends because, ultimately, this is a good thing, you never see this far, and falsely believe that there are genuine reasons of loyalty. And this might be the best state of affairs.

Why is this objectionable? Williams[39] objects that, if self-effacing consequentialism were

true, then nobody ought to believe it. Self-effacing consequentialism tells us to see certain considerations as practical reasons when they are, by its lights, not. And it tells us to deny that certain factors are practical reasons when, by its lights, they are. We take it that moral reasoning is a species of practical reasoning concerning moral rectitude. And we have a picture of practical reasoning according to which practical reasoning involves determining which considerations are practical reasons. A self-effacing consequentialism is inconsistent with this conception in the sense that, by the lights of the theory, things go best only if we remain ignorant of (many of) our fundamental practical reasons. We are debarred from being robust practical reasoners.

Is this coherent? Self-effacing consequentialisms give practical grounds for our having false beliefs about our practical reasons (grounds of which we must, of course, remain ignorant lest we lose the false beliefs). But there are cases where this seems quite coherent. If you knew about the lurking lion, this would cause you to sweat, thereby enabling it to smell you. Thus there is a practical ground for remaining ignorant. You have a reason for believing in the absence of lions that has no bearing on the claim that they are absent. Self-effacing consequentialisms simply embrace such beneficial ignorance on a larger scale. But while this may be coherent, there are theoretical costs. In the case of the lurking lion there is a backdrop of practical ends (not being devoured being a prominent one) of which the agent is aware and towards the achievement of which the ignorance contributes. But self-effacing consequentialisms require general false beliefs about practical ends, and it seems less plausible that adept practical reasoners could be *generally* mistaken about their practical *ends.*

Rossian deontology does not have to bear such theoretical costs. It gives a straightforward account of our obligations toward our friends: there are basic agent-relative reasons to favor friends, some of which are of sufficient exigency in certain circumstances to constitute obligations.

But are these special obligations unacceptably partial? No. In our view, we show partiality in allocating goods only if we give the claims of one person or group more weight than we are *warranted* in doing (partiality is an irreducibly normative notion). The different theories all respect the need to be impartial, but offer competing conceptions of what features are relevant in assessing whether one person's claim is weightier than another's. If, as welfare consequentialism maintains, people's claims are proportional to the effect on their interests, then we would show partiality if we gave greater weight to the interests of some particular person(s). Unlike consequentialism, Kantianism and contractualism put obligations to other people centre-stage. But they treat persons impartially by making no distinction among persons: each owes the same to every other, simply in virtue of all being persons. Thus Korsgaard, a leading Kantian, (1996b, 126-128) and Scanlon (1998, 160-162), while admitting that friendship has many of the structural features of morality, deny that it is a basis of *moral* obligation. If, however, as we are claiming, we have moral ties to friends and family we are not showing partiality to them merely in virtue of putting their interests above those of others. We act partially only if, like the clannist, we give *undue* weight to those interests, more than is warranted by the relationship.

Must the Rossian deny that friendship is valuable? No. Friendship is valuable and there are reasons to promote friendship in general, but your reasons to favor your own friends do not derive from these. And that you could better promote friendship by abandoning your friends would not furnish you with sufficient reason to do so.

Nor does the Rossian see duties of special obligation as inexplicable. But the explanation does not appeal to the value. Rather, reasons of friendship, such as reasons to be loyal, cannot be derived from anything more fundamental, just as for the consequentialist the proposition that, say, human welfare is a good, cannot be derived. But basic reasons of friendship, like fundamental propositions about what is valuable, can be explained by

incorporating them into a well-articulated account of morality. The choice is ultimately between differing overall such accounts. We have suggested that consequentialist accounts fare worse than Rossian accounts when it comes to special relationships.

We are not wholly in agreement with Ross, however. On his view there is a duty to be beneficent provided that one is not subject, in the circumstances, to a duty more stringent. Thus all reasons to favor friends must be moral, lest they carry no weight against the duty to be beneficent. On our view, there is a duty to be beneficent on occasion, but it is not pervasive. Thus there is room for non-moral reasons to favor friends. More generally, Ross's duty of beneficence rules out options – our next topic.

7.2   *Options*

Suppose pleasure is a good. And suppose that on some occasion Al receives some pleasure and Betty experiences pain, with the result that the net amount of pleasure in the world is increased. On one view, perhaps, the amount of good in the world is increased, but Al does not receive a benefit that Betty is denied. Benefits cannot go to individuals, and thus the debate about distribution cannot get off the ground. We set aside this position, however, on the grounds that most consequentialists would find it as counterintuitive as we do (one of the initial attractions of consequentialism is the thought that we should make as many lives as possible go as well as possible; but to make a life go well is to provide its liver with benefits).

On the assumption that benefits can go to individuals, we can ask whether each person has special reason to pursue her own benefit. We think she does. But on SC, it seems, she does not: she only has reason to pursue her own benefit insofar as its pursuit will contribute to her maximization of the general good, and this is not a special reason that she has and others lack.

SC, then, leaves no room for the pursuit of, say, personal projects, unless their pursuit maximizes the good. However, there seems to be a rationale for their pursuit even in the face of

their sub-optimality. We each have special 'personal reason' to care about our own interests and concerns just because they are ours. These agent-relative personal reasons arise because each of us has our own point of view. I have a personal reason to care about my pain that I cannot have to care about yours, namely that it is mine. This does not mean that I have no reason to care about your pain, nor does it commit me to denying that pain is equally bad whoever has it. Personal reasons, then, (to state matters in consequentialist terms) give each agent moral permission – i.e., the option – not to maximize the good when the cost to *her* would be significant[40]. An agent is allowed, in determining what she is morally required to do, to accord greater weight to the cost borne by her than is warranted by its impersonal disvalue.

How can this be? Since I am a creature with a personal point of view, who has personal reasons, a morality that required me to transcend that point of view and think of the world as if I had no particular place in it, would not merely be unreasonably demanding, it would deny all moral significance to the fact that my life is, in a sense, all I have. There has, therefore, to be some balance between the demands that the needs of others put on us and our right to live our own lives. Determining where that balance lies is notoriously difficult. No doubt we are inclined to suppose that morality is less demanding than it is. But this does not entail that there is no balance to be struck.

SC cannot accommodate personal reasons if it sees all reasons as stemming from agent-neutral value – it can then at best hold that, since people care disproportionately about their own good (though they have no reason to), their failure to maximize good is understandable when the cost to them is high. But to understand is not to justify.

A weaker thesis is that all *moral* reasons are agent-neutral. SC might allow that personal reasons can mount up to give us sufficient reason not to do the right act. That concession, however, will not give us options, for we will not have *moral* permission to bring about less good than we could. Sophisticated C may appear to give us such moral permission since it tells

agents that they *should* develop dispositions that may lead them to do wrong. However the justification for that advice lies not in the moral significance of personal reasons, but in the claim that giving some priority to our own concerns will bring about better long-term results than if we try to act rightly on each occasion.

AC, then, cannot accommodate the moral significance of personal reasons. Given the AC framework, this denies room for supererogation (acting beyond the call of duty): the person who bears great personal cost in maximizing the good, although admirable in the extreme, would be doing something morally wrong if she did otherwise.

RC does better. An agent who follows the rules does not act wrongly: she does enough good – it is meritorious but not required to do more. The presence of this personal space, however, stems from impersonal costs: we are psychologically resistant to making significant sacrifices, and this makes it too expensive to inculcate a more demanding rule. But this resistance is a regrettable flaw, not a mark of personal reasons. Even if RC concedes that the resistance has a rational basis, it is still committed to denying the *moral* significance of personal reasons at the fundamental level. They matter morally only because of the cost of training people to ignore them.

On one deontological view, in deciding what to do the virtuous agent balances the good to be achieved, and for whom, against her cost. Although her cost makes no moral claim on her, personal reasons are nevertheless morally significant because she is morally permitted not to bear the cost if it is disproportionately heavy. On this picture, the agent may not be morally required to satisfy the weightiest moral claim – but if, under these circumstances, she does, she is praiseworthy not only for doing good but also for doing it supererogatorily.

Deontology also permits us to choose how to exercise our beneficence, if no other obligations are in play. People are free to take up causes dear to their hearts, without this being part of a strategy for maximizing the good.

Supererogatory acts do not require either saintly or heroic qualities. Small sacrifices can be supererogatory. Many quite trivial acts of kindness are like this. Whether helping others is supererogatory as opposed to morally required depends on, among other things, the relative size of the benefit and the sacrifice, and the relation in which you stand to the beneficiary.

Most contemporary defenses of supererogation rest on the claim that agents cannot be required to do good if the cost would be disproportionately great. Some believe that this concedes too much to consequentialism from the start. It implies that we would have a duty to do as much good as we could, if it were cost free. Traditional theology[41] has denied this. God has a duty, perhaps, to grant us lives worth living, but not to grant anything more, even though to Him it is cost free.

7.3   *Constraints*

Constraints, though often regarded as the most distinctive feature of deontology, seem hard to justify. Consider an absolute constraint (C) against (intentionally) killing an innocent person. Suppose Anne and several other innocents are about to be shot by Bert, but he agrees to let the others go if you shoot Anne. (C) forbids you to do it. Yet, as Scheffler[42] points out, this appears inexplicable: Anne is going to be shot, but at least you can prevent the other shootings. The standard objection to constraints is that they forbid their own violation *even to minimize such violation.* Another difficult case is one in which violation will result in some other good that outweighs the bad of the violation. The general SC complaint against constraints, be they absolute or *pro tanto*, is that they can *forbid* you to do good.

There are proscriptions acceptable to both SC and us – but they are not constraints. Recall Scanlon's remark that 'the fact that a course of action can be foreseen to lead to someone's death is normally a conclusive reason against it; [and] the fact that someone's death would be to my personal advantage is no justification for aiming at it' (1998, 199). This implies that there is an absolute prohibition against killing purely for personal gain. That is, there is an

absolute prohibition against killing another person when one's only motivation is personal gain, and when, in fact, there are no (other) reasons to kill. (This is not quite our final position on this issue; but laying out our final position would take us too far afield from the concerns of this essay.) Such a ban is acceptable to us since one could only violate it for reason of one's own personal gain, which is no reason to kill someone. And since killing someone purely for personal gain does not increase the good, one cannot violate the ban in order to do so, hence it is acceptable to SC. Constraints, on the other hand, are proscriptions that admit the possibility of, and forbid, their own violation to good effect.

Another important feature of constraints as understood by traditional deontology is that they are underivative. RC, for example, incorporates proscriptions, but these are 'justified by their role in promoting agent-neutral value' (Hooker 2000, 110).

We have defended special obligations and options by contending that, in addition to the amount of good we do, positional facts – that the good would accrue to *my* friend or to *me* – are also morally relevant. Constraints, however, cannot be similarly defended. What justifies constraints? That their violation is bad is no answer: for then how could it be forbidden for someone to violate a constraint in order to prevent worse actions by others? The strategy of introducing morally relevant positional facts does not help. Constraints single out no group on the basis of my relationship to its members, thus they cannot rest on my being more closely related to some than others. Hence the only positional possibility is to claim that my violating a constraint, even to prevent worse actions by others, is a bad *for me.* But although such a violation may matter to me, since I have a perfectly understandable, if perhaps not always commendable, reluctance to get my hands dirty, this, at best, *might* ground a *permission* not to violate the constraint under these conditions. It cannot ground a *requirement* not to do so. It is implausible to suppose that a constraint violation is wrong because we have an aversion to it; and to say that we have an aversion to it because it is wrong does nothing to explain why it is

wrong.

Constraints embody as fundamental the fact of my agency. I should not, claims the defender of a constraint against killing, aim directly at someone's death myself, even when my ultimate goal in doing so is to thwart the similar killing aims of several others. But why, SC asks, should this disposition of the will matter, from the moral point of view, when my ultimate goal is good (fewer killings)? One response is simply to claim that, in addition to doing good, I should respect constraints, and the latter takes priority. Constraints are fundamental so we should not expect to find a deeper justification for them. Their being fundamental does not, however, preclude our defending constraints, as we did with special obligations and options, by explicating their nature in ways that make their force clearer. The problem is that we can find no explication of constraints that dispels their air of irrationality in light of cases such as that of Anne and Bert.

Many deontological attempts to explain why agency matters, for example, seem to presuppose the very point at issue[43]. Thus it is said (in a Kantian vein) that persons deserve respect in view of their unique importance as rational moral agents. But why does such respect forbid you to harm others rather than requiring you to minimize harm? It may be said that, just as we owe particular duties to others in view of our special relationships with them, so we owe to everyone else a duty not to harm them because of our general relationship with them. But what is that relationship? Perhaps that of being fellow humans, or fellow persons. Whatever the answer, the problem remains: why does our standing in that relationship to all ground a *constraint* against harming them as opposed to a duty to minimize harm? Similarly, Natural Law theorists move from the claim that there are certain basic values, including life, to the claim that we should never act directly against a basic value, even in seeking to protect that value elsewhere. But how is that move to be justified?

Another Kantian line is to claim that we cannot be responsible for another's will. Thus even though you may prevent others, say, killing innocents by killing one yourself, you would not thereby have prevented those others harboring evil intent. And you are neither responsible for those intentions nor for their fulfillment. Whilst we have some sympathy with this line, nevertheless standing idly by when you could achieve an overall net saving of human life on the grounds that its loss is not your responsibility suggests a squeamish desire to keep *your* hands clean.

Some defenders of constraints[44] have complained that, in seeing constraints as *agent-*relative, recent attempts to ground constraints have wrongly focused on agency. Rather, they claim, we should focus on a *patient-*centered justification – on what it is about innocents that entails the existence of constraints against harming them. But this does not seem to help. By the nature of innocence, innocents do not deserve to be harmed, so that, ideally, we should not harm them. But what are we to do in our non-ideal world in which innocents are under threat?

Quinn, Kamm and Nagel[45] suggest that it is impermissible to kill the innocent (even to prevent further such killings) because the world is better for having such an innocence-respecting constraint. Fewer innocents might be killed in a world where such a constraint is lacking – it being permissible to kill innocents there to minimize the number of innocents killed. But, the claim runs, the reduction in the amount of killing in this latter world notwithstanding, our constrained world is better because innocent life is more valuable here due to the impermissibility of its sacrifice. Even if this is correct, however, why is the constrained world actual? It is not valid to argue from betterness to truth (for instance our longevity does not follow from the premiss that it would be better if we all lived longer). Nagel sees the difficulty, but thinks that such reasoning 'may have a place in ethical theory, where its conclusion is not factual but moral' (1995, 92). In our view, however, this is to justify constraints at the expense of standards of justification.

If we deny the existence of constraints, however, won't we have to abandon many of our intuitions? No. First, as we have seen, certain proscriptions are acceptable: they are bans on doing things *for inappropriate reasons*. Second, many intuitions that appear to be based on constraints are actually based on other features. We may, for example, think it wrong to take $10 from one person in order to enrich another by $20. But our grounds for that (depending on the circumstances) might be that the harm of taking $10 *honestly possessed* outweighs the benefit of bestowing $20 *unearned*. Or we may think it wrong to do considerable harm to one person in order to prevent small harms to a large number. But that may be because harms are lexically ordered. Finally, acts that are considered by traditional deontologists to be violations of serious constraints will increase good only in dire circumstances.

We are tentatively proposing, then, a morality devoid of constraints (as traditionally understood) but incorporating duties of special obligation and options. Hence we must answer arguments to the effect that constraints are required to 'protect' our relations with our friends and family, and our option, say, to limit our charitable donations.[46]

Suppose I am proposing to buy a house, and you note, correctly, that I could do more good by buying a cheaper house and donating the savings to Oxfam. I reply that doing so would be supererogatory. You then argue that you are permitted to maximize the good, and will set about frustrating my attempt to purchase the more expensive house. Or suppose that I refuse to be disloyal to a friend. You see that my loyalty here is not maximizing the good, and set about undermining it.

We have argued that each of us has reasons, say, to favor friends and pursue personal projects. I can acknowledge that your friendships are valuable, and that at least some of your personal projects are worthwhile. Thus insofar as I have reason to promote the good, I have reason to promote your friendships and your pursuit of those projects of yours that are

worthwhile. But perhaps I can do more good by interfering with your friendships and your projects. What reason is there to desist, unless there are constraints against my so interfering?

However, the standard difficulty with constraints applies here also. Suppose there is a constraint against interfering with another's pursuit of his friendships. Then I am forbidden to interfere with your friendships even to prevent much greater such interference by someone else. It seems to us more plausible to maintain that non-interference is itself valuable. So that, in addition to the general value of your friendships and worthwhile pursuits, there is an additional reason for my non-interference. This even affords some protection for your trivial pursuits. But in no case is the protection absolute, of course: there can be greater values at stake, for example.

Kagan[47] mounts a different argument against Scheffler's attempt to have options without constraints. Suppose there is an option to forego saving the life of a stranger if it will cost you $10,000. Then why are you not permitted to kill a stranger to gain $10,000?[48] The former permission seems reasonable, but the latter is not. Thus, it appears, any moral system that incorporates the first option must incorporate a constraint against killing.

If one notes that it is not a constraint against killing simpliciter that is required, but an acceptable ban against killing for personal gain, one must face the issue of why there is not also an acceptable ban against refusing to save the life of a stranger for selfish reasons. (Note that there is an acceptable ban against doing nothing to save the lives of strangers for selfish reasons – i.e., there is a requirement to do something for charity.) Kagan's opponent might argue that refusing to save here can be outweighed by the relevant personal reasons whereas the proposed killing cannot. There are at least three potentially relevant differences (whether they are actually relevant depends upon the circumstances) that support this. First, killing may be in itself worse than not saving. Second, there may be a morally relevant difference between donating a sum of significance to you and not gaining that sum. Both these claims of moral difference are hotly contested,[49] but our position is that they can be relevant on occasion.

The third difference is that, in the case of the killing, you would be solely responsible, whereas in the case of not saving you would typically share responsibility with all the others who could have contributed but didn't. Where you would not share responsibility, and there are no other relevant differences between the killing and the refusal to save, we suspect that there is no option not to save. Suppose, say, that, by no fault of yours, your life savings are about to be burned at the local incinerator, and you are rushing to retrieve the money. If you delay, you know you will lose it. In case (1) you must stop if you are not to kill a stranger lying unconscious in the road. In case (2) you are the only person capable of saving a stranger from being killed by an oncoming vehicle, but you must stop if you are to do it. Are we committed to saying that you are morally required to stop in case (1) but not case (2)? No. Our view is that you must stop in both cases. The fact that you are the *only* person capable of saving this life is a crucial morally relevant consideration in case (2): it is one of the factors that here renders irrelevant the difference between harming and failing to prevent harm. But to conclude from this that such a difference is never relevant is to over-generalize.

We are not, then, in Kagan-like examples, generally committed to the permissibility of harming to gain a benefit *b* in cases that are analogous to cases in which it is permissible to fail to prevent a similar harm in order to save oneself from a cost of *b*, because there may be relevant differences between harming and failing to prevent harm, or between gaining a benefit and saving oneself from a cost, that disrupt the purported analogy.

A theory with special obligations and options, but without constraints, is still deontology: agent-relative considerations are underivatively relevant, and agents are forbidden in certain circumstances to maximize the good. On Ross's view there are, of course, no *absolute* constraints. But Ross may nevertheless be interpreted as claiming that there is a constraint against harming. If so, we tentatively disagree with him. On our alternative, personal reasons

and those stemming from special relationships carry significant weight on the moral scales, but mere reluctance to harm in the service of good that outweighs that harm carries none.

7.4   *Virtue Ethics*

Consequentialism and deontology do not exhaust the options for a moral theory. Virtue ethics rejects what it sees as serious defects in both approaches, particularly their focus on the deontic status of acts, and their belief that moral theory should formulate precise moral principles from which to read off conclusions about what to do. Good moral judgment requires sensitivity, experience, and discernment, rather than slavish adherence to pre-determined rules. The virtues are valuable in their own right, and not just as a guarantor of reliably choosing the right act. And only those who possess them can discern what is morally salient in any particular situation. The moral emotions play a crucial role not only in determining how one should act but also in motivating the agent; the sense of duty is to be harnessed only when better motives fail. Moral philosophy should focus more, therefore, on what kind of person it is best to be, rather than on what principles we should invoke to solve artificially constructed moral dilemmas.

There are two views concerning the purport of virtue ethics. On the less radical, virtue ethics is proposed as a welcome corrective to various distortions that have afflicted many versions of both deontology and consequentialism. On the more radical, it is put forward as an alternative theory in its own right. With the less radical approach we are entirely in sympathy. The deontological theory we favor, which is broadly Rossian in spirit, takes the above points on board. We see overly principled approaches as distorting moral thinking by downplaying the need for judgment and imagination in discerning, in a particular case, which features are relevant, how they interact with each other, and what weight should be given to each. And it has often been pointed out that deontology and virtue ethics make common cause against consequentialism[50]. But we are skeptical of the more radical approach, which maintains that the right is metaphysically dependent upon the judgments of the virtuous. In difficult cases we may

have no *epistemic* access to which act is right other than via the judgment of virtuous agents. But the virtuous agent judges an act right because it is right, not the other way around. Otherwise, on what does she base her judgment? It must be responsive to reasons, and those reasons, if she is appropriately sensitive, lead her to the truth. The latter is there for her to find; she does not construct it. Virtue ethics is thus best seen as a crucial part of the best deontological theory.

Bibliography

Adams, R. (1976) 'Motive Utilitarianism', *Journal of Philosophy* 73, 467-481.

Arneson, R. (2003) 'Consequentialism vs. Special-Ties Partiality', *Monist* 86, 382-401.

Audi, R. (1996) 'Intuitionism, Pluralism, and the Foundations of Ethics', in Sinnott-Armstrong, and Timmons (1996), 101-136.

Blackburn, S. (1999) 'Am I right?' *New York Times* February 21.

Brink, D. (forthcoming) 'Some forms and limits of consequentialism' in Copp (forthcoming).

Brook, R. (1991) 'Agency and Morality', *Journal of Philosophy* 88, 190-212

Crisp, R. and Slote, M. (1997) *Virtue Ethics*. Oxford: Oxford University Press.

Copp, D. (1979) 'The Iterated-Utilitarianism of J.S. Mill', *Canadian Journal of Philosophy*, Suppl Vol V, 75-98.

Copp, D. (1995) *Morality, Normativity and Society*. Oxford: Oxford University Press.

Copp, D. (1997) 'Does moral theory need the concept of society?', *Analyse & Kritik* 19, 189-212.

Copp, D. (ed.) (forthcoming) *The Oxford Handbook on Moral Theory*. Oxford: Oxford University Press.

Dancy, J. (1983) 'Ethical Particularism and Morally Relevant Properties', *Mind* 92, 530-47.

Dancy, J. (1993a) *Moral Reasons*. Oxford: Blackwell.

Dancy, J. (1993b) 'An Ethic of *Prima Facie* Duties', in Singer (1993), 230-240.

Darwall, S. (1986) 'Agent-Centered Restrictions from the Inside Out', *Philosophical Studies* 50, 291-319.

Darwall, S. (forthcoming) 'Morality and Practical Reason: A Kantian Approach' in Copp (forthcoming).

Driver, J. (2001) *Uneasy Virtue*. Cambridge: Cambridge University Press.

Hare, R. M. (1963) *Freedom and Reason*. Oxford: Clarendon Press.

Hare, R. M. (1981) *Moral Thinking: Its Levels, Method, and Point.* Oxford: Clarendon Press.

Hill, T. (forthcoming) 'Kantian Normative Ethics' in Copp (forthcoming).

Hooker, B. (2000) *Ideal Code, Real World*. Oxford: Clarendon Press.

Hooker, B. (1994) 'Is Rule-Consequentialism a Rubber Duck?' *Analysis* 54, 62-7.

Hooker, B. and Little, M. (eds.) (2000) *Moral Particularism.* Oxford: Clarendon Press.

Howard-Snyder, F. (1993) 'Rule Consequentialism is a Rubber Duck', *American Philosophical Quarterly* 30, 271-8.

Hurka, T. (2001) *Virtue, Vice and Value.* New York: Oxford University Press.

Hursthouse, R. (1999) *On Virtue Ethics.* Oxford: Oxford University Press.

Kagan, S. (1984) 'Does Consequentialism Demand Too Much?', *Philosophy and Public Affairs* 13, 239-54.

Kagan, S. (1989) *The Limits of Morality.* Oxford: Clarendon Press.

Kamm, F. (2000) 'Nonconsequentialism', in Lafollette (2000), 205-226.

Kamm, F. (1989) 'Harming some to save others', *Philosophical Studies* 57, 227-260.

Kamm, F. (1992) 'Non-consequentialism, the person as an end-in-itself, and the significance of status', *Philosophy and Public Affairs* 21, 381-389

Kant, I. (1993) *Grounding of the Metaphysics of Morals* (originally published 1785) trans. J. Ellington. Indianapolis: Hackett.

Korsgaard, C. M. (1996a) *Creating the Kingdom of Ends.* Cambridge: Cambridge University Press.

Korsgaard, C. M. (1996b) *The Sources of Normativity.* Cambridge: Cambridge University Press, 1996b.

Lafollette, H. (2000) *The Blackwell Guide to Ethical Theory.* Oxford: Blackwell.

McGinn, C. (19990 'Reasons and unreasons', *The New Republic*, May 24, 34-38.

McNaughton, D. (1988) *Moral Vision.* Oxford: Blackwell.

McNaughton, D. (1996) 'An Unconnected Heap of Duties?', *Philosophical Quarterly* 46, 433-447.

McNaughton, D and Rawling, P. (1991) 'Agent-Relativity and the Doing-Happening Distinction', *Philosophical Studies* 63, 167-185.

McNaughton, D. and Rawling, P. (1993) 'Deontology and Agency', *The Monist* 76, 81-100.

McNaughton, D. and Rawling, P. (1998) 'On defending deontology', *Ratio* 11, 37-54.

McNaughton, D. and Rawling, P. (2000) 'Unprincipled Ethics' in Hooker and Little (2000), 256-

275.

McNaughton, D. and Rawling, P. (2003a) 'Can Scanlon avoid redundancy by passing the buck?', *Analysis* 63, 328-331.

McNaughton, D. and Rawling, P. (2003b) 'Naturalism and Normativity', *Proceedings of the Aristotelian Society,* Supp. Vol. 77, 23-45.

Nagel, T. (1986) *The View from Nowhere.* Oxford: Oxford University Press.

Nagel, T. (1995) 'Personal rights and public space', *Philosophy & Public Affairs* 24, 83-107.

Parfit, D. (1987) *Reasons and Persons* (rev. edn.). Oxford: Clarendon Press.

Quinn, W. (1993) *Morality and Action.* Cambridge: Cambridge University Press.

Rachels, J. (ed.) (1998) *Ethical Theory.* Oxford: Oxford University Press.

Railton, P. (1988) 'Alienation, Consequentialism, and the Demands of Morality', in Scheffler 1988, 93-133. (Originally appeared in *Philosophy and Public Affairs* 1984, 13).

Regan, D. (1980) *Utilitarianism and Cooperation.* Oxford: Clarendon Press.

Ridge, M. (2001) 'Saving Scanlon: Contractualism and Agent-relativity', *The Journal of Political Philosophy* 9, 472-81.

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

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

Scheffler, S. (1994) *The Rejection of Consequentialism*, rev. edn. Oxford: Clarendon Press.

Scheffler, S. (ed.) (1988) *Consequentialism and its Critics.* Oxford: Oxford University Press.

Singer, P. (ed.) (1993) *A Companion to Ethics.* Oxford: Blackwell.

Sinnott-Armstrong, W. and Timmons, M. (eds.) (1996) *Moral Knowledge: New Readings in Moral Epistemology.* Oxford: Oxford University Press.

Smart, J.J.C. (1998) 'Extreme and Restricted Utilitarianism' in Rachels (1998), 286-297.

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

Stratton-Lake, P. (2003) 'Scanlon's contractualism and the redundancy objection', *Analysis* 63, 70-76.

[1] For a careful statement of this position, see Arneson 2003 p. 382.

[2] For this terminology, see e.g. Nagel 1986, ch 9, and Kagan 1989.

[3] Note that some deontologists, such as W. D. Ross (1930), reject options, maintaining that we ought to do as much good as we can, limited only by duties of special obligation and constraints.

[4] This is not true of all forms of consequentialism: see below.

[5] Nagel 1986, Darwall 1986, Parfit 1987, Scheffler 1994, McNaughton & Rawling, 1991.

[6] But see McNaughton & Rawling 1991 for discussion of some problems for this approach.

[7] Parfit 1987, 27.

[8] Parfit (1987, 40) notes that Sidgwick coined this phrase.

[9] See for instance Kagan 1989, Smart 1988, 288-89.

[10] E.g., Brink (forthcoming).

[11] Parfit sees such actions as cases of 'blameless wrongdoing':

If we have one of the best possible sets of motives, we shall sometimes knowingly act

wrongly according to our own theory [viz. Sophisticated C]. But, given the special reason why we are acting wrongly, we need not regard ourselves, when so acting, as morally bad. We can believe these to be cases of *blameless wrongdoing.* We can believe this because we are acting on a set of motives that it would be wrong for us to cause ourselves to lose. (Parfit, 1987, 49, italics his.)

[12] This issue perhaps does not arise if RC is seen as merely laying down a criterion of wrong action in terms of rules, where these rules need never have been formulated or internalized. But we shall assume that 'the rules are to be *public*' (Hooker 2000, 85, italics his), and are 'inculcated … by family, teachers, and the broader culture' (2000, 79).

[13] cf. Adams 1976

[14] Copp's Mill (1979) may be another case in point. On this interpretation of Mill:

> S's doing A would be wrong if, and only if, (a) there is a maximal alternative to S's doing A, and (b) it would be maximally expedient that if S did A, S would feel regret for this to some degree (Copp 1979, 84).

Thus if an act is wrong it is not best, but its failure to be best does not entail its wrongness. And, we suspect, Copp's Mill would not advocate that agents employ thoughts about clauses (a), (b) above in deliberating about what to do in every case.

[15] The best motives, it is claimed, need not produce the best acts – we set aside here discussion of the coherence of this claim (see Parfit, 1987, 31-35, 37-40 for discussion of this sort of issue); a similar worry about coherence might also arise in the case of Sophisticated C.

[16] RC incorporates options, but their justification is not agent-relative – it appeals, rather, to the cost of inculcating a rule that requires consistent significant sacrifice (and, perhaps, to the value of people being permitted to pursue personal projects?)

[17] For a dissenting voice on the issue of whether RC is genuinely consequentialist, see Howard-Snyder 1993. For Hooker's reply, see Hooker 1994.

[18] For a full exposition of Copp's view, see Copp 1995.

[19] This last is, we think, agent-relative for Ross: he does not appear to countenance doing harm oneself in order to minimize the total amount of harm. (For a rebuttal of the complaint that this is an unsystematic list see McNaughton 1996.)

[20] Audi (1996) points out that someone can know a self-evident truth without knowing that it is self-evident. Moreover, from the fact that a claim needs no justification it does not follow that it has none.

[21] Ross's equation of the unconditional with the underivative might explain why he sees *prima facie* duties as on a par with simple mathematical truths. In the mathematical and the ethical cases one can see that certain truths are basic. Thus, thinks Ross, like the mathematical truth, the ethical truth is unconditional. Unfortunately for Ross, however, the ethical case, unlike the mathematical, may be conditional.

[22] A view stoutly maintained by R. M. Hare. See, for example, 1963, ch 2, esp. p. 7.

[23] See McNaughton and Rawling 2000.

[24] Since Kant is so adequately covered elsewhere in this volume – see Hill and Darwall (forthcoming) – we shall be brief.

[25] Of one version of the CI test – contradiction in the world. There is also contradiction in willing.

[26] See Hill (forthcoming).

[27] This interpretation of Kant is disputed (see Hill (forthcoming), esp. footnote 2.) Kant does not appear to take this line in the *Grounding of the Metaphysics of Morals*. In discussing the second formulation of the categorical imperative which requires us to treat others always as ends in themselves, and never simply as means, he writes 'concerning meritorious duty to others' that merely refraining from impairing the happiness of others is not sufficient. For 'this, after all, would harmonize only negatively and not positively with humanity as an end in itself, if everyone does not also strive, *as much as he can*, to further the ends of others. For the ends of any subject who is an end in himself must *as far as possible* be my ends also, if that conception of an end in itself is to have its full effect in me.' (1993, 37, our emphasis)

[28] Korsgaard suggests that we can read Kant in a slightly less rigoristic manner in 'The Right to Lie: Kant on Dealing with Evil' in Korsgaard 1996a, 135-158. See also 'Two Arguments Against Lying' in the same volume, 335-362.

---

[29] See his discussion of constraints, Scanlon 1998, 81-86.

[30] We are modelling this discussion on Scanlon's discussion of his principle F of fidelity on p.304, where he speaks of reasons to want a principle. Scanlon's principles lay out what is right and wrong; and the fact that some act would be, say, wrong, is itself a reason not to do it, on Scanlon's account. Thus, on this account it seems, we have reason to want there to be certain further reasons, and this is relevant in determining whether there are such further reasons. On our view, by contrast, you might have reason to want, say, that I φ because my doing so will benefit you; but whether the fact that my act will benefit you is a reason for me to do it is not dependent on whether you want it to be a reason.

[31] There has been much discussion of what has come to be called 'the redundancy objection' first raised by Pettit (1993, 302). See, for example, Blackburn 1999, McGinn 1999, Ridge 2001, Stratton-Lake 2003, and McNaughton and Rawling 2003a.

[32] Scanlon (1998, 7) thinks that 'it is not clear that morality in the broader sense is a single subject that has a similar unity.'

[33] On this last point we agree with Scanlon; see also McNaughton and Rawling (2003b).

[34] MC and Sophisticated C might respond by arguing that, given the tendencies many of us have to resentment and other forms of psychological resistance to altruism, the best disposition *for such individuals* might not be to make continuous significant sacrifice on behalf of those in poverty. RC, however, does not rely in this way on contingent empirical facts about differences

in individual psychology. RC might maintain that the right act accords with those motives which are such that *if everyone had them*, things would go best.

[35] Smart (1998, 292) accuses RC of 'superstitious rule-worship'.

[36] Hooker, 2000, endeavors to circumvent this objection by resting the theory on value *plus impartiality*.

[37] Copp's Mill does not escape this sort of objection: according to him it is not wrong to favor friends at the expense of failing to maximize the good because it would not be maximally expedient to feel regret for doing so. But surely, rather, it is permissible to favor friends at the expense of failing to maximize the good because they are your friends.

[38] Parfit, 1987, 40-43.

[39] in Smart and Williams, 1973, 135.

[40] For an extended defense of this approach, see Scheffler *passim*.

[41] See, e.g., Adams 1972.

[42] Scheffler 1994.

[43] A point made with force by Scheffler 1994, ch 4.

[44] See for instance Brook 1991, Kamm 2000. For discussion see McNaughton and Rawling 1993.

[45] Kamm 1989 & 1992, Quinn 1993, Nagel 1995. For a longer discussion of this point see McNaughton and Rawling 1998.

[46] Of course, the defender of constraints might buy the arguments below to the effect that special relationships and options entail constraints, and see these arguments as arguments for constraints.

[47] See Kagan 1984. Scheffler responds in Scheffler 1994, 167-192.

[48] Kagan's (1984) original example contrasts saving a stranger with killing one's uncle. We have modified it to avoid considerations of special relationships.

[49] For a sustained and vigorous attack see Kagan 1989.

[50] Recent writers (e.g. Hurka 2001, Driver 2001) have argued that consequentialism can incorporate many of the claims of virtue ethics in its less radical version. For a helpful collection of articles on virtue ethics see Crisp and Slote 1997. For an exposition and defense, see Hursthouse 1999.