<center>Value and Agent-Relative Reasons</center>

In recent years the distinction between agent-relative and agent-neutral reasons has been taken by many to play a key role in distinguishing deontology from consequentialism.[1] It is central to all universalist consequentialist theories that value is determined impersonally; the real value of any state of affairs does not depend on the point of view of the agent. No reference, therefore, to the agent or to her position in the world need enter into a consequentialist understanding of what makes an action right or wrong or morally permissible. Consequentialism thus provides an agent-neutral account of both the right and the good.

Recent opposition to consequentialism has focussed on two aspects of common-sense morality -- constraints and options[2] -- which do not appear to be adequately captured by an agent-neutral understanding of right and wrong, but which are typically given a central place in deontological theories. In both cases reference to the agent enters essentially into an understanding of the moral quality of the action. In the case of a constraint, what matters is that the agent not perform the act in question. In the case of an option not to maximise the good, what matters is not merely the value of, say, my project, but the fact that it is mine. Both constraints and options, then, are agent-relative, but in other ways are distinct. In each case the agent has a reason for acting which others do not share, but the reason is a requirement only in the case of constraints.

The notion of a reason which others do not share is ambiguous in an important respect. Parfit[3] distinguishes between aims which "are held in common" and those which are not. The former

---

[1]. S. Scheffler, The Rejection of Consequentialism, Oxford, 1982; D. Parfit, Reasons and Persons, Oxford, 1984, p. 27; 1982; T. Nagel, The View from Nowhere, Oxford, 1986, chs 8,9; S. Darwall, 'Agent-Centred Restrictions from the Inside Out', Philosophical Studies 50, 1986, 291-319; P. Pettit, 'Universalizability without Utilitarianism', Mind, 96, 1987, 74-82; D. McNaughton and P. Rawling, 'Agent-Relativity and the Doing-Happening Distinction', Philosophical Studies, 63, 1991, 167-85; D. McNaughton and P. Rawling, 'Honouring and Promoting Values', Ethics, 102, 1992, 835-43; D. McNaughton and P. Rawling, 'Deontology and Agency', The Monist, 76, 1993, 81-100; J. Dancy, Moral Reasons, Oxford, 1993, chs. 10,11.

[2]. The terminology we use here is that of Shelly Kagan (The Limits of Morality, Oxford, 1989). This terminology is neither universal (Scheffler, for example, calls the former 'agent-centred restrictions' and the latter 'agent-centred prerogatives') nor uncontentious, since deontologists can rightly complain that the notion of an option is one that makes sense only from within a fundamentally consequentialist framework. However, the terms have the advantages of familiarity and brevity.

[3]. In Reasons and Persons -- see, e.g., p.27 and p.104

aims he dubs agent-neutral, the latter agent-relative. Following Parfit's lead, one sense in which agent-relative reasons are not shared is that they furnish each of us with a different aim. However, some authors (notably Thomas Nagel and Christine Korsgaard[4]) have a tendency to think in terms of agent-relative and agent-neutral <u>value</u>: from this perspective, agent-relative reasons are not shared because they derive from value-for-the-agent, whereas agent-neutral reasons are shared since they derive from "impersonal" (Nagel, p.153) value.

Characterising deontology in terms of agent-relative value gives rise to various purported difficulties for the position, which, we shall argue, are spurious -- defining deontology in terms of agent-relativity appropriately construed gives rise to none of them. Korsgaard (see, e.g., pp.47-48) objects to defining deontological reasons as agent-relative or "subjectively normative" (p.47). However, we shall argue that, as we define the notion of an agent-relative reason, those agent-relative reasons corresponding to constraints are not "subjectively normative" because they are not tied to agent-relative value. Thus, whereas Korsgaard sees the source of various spurious complaints against deontology as located within what she regards as the mistaken association of deontology with agent-relativity, we see these complaints as arising from the inappropriate equation of agent-relativity with agent-relative value. We do agree with Korsgaard that agent-relative reasons are not (generally, at least) shared, but, for constraints, this is only in the Parfitian sense.

Two of the complaints to which we refer are raised by Korsgaard herself. First, since agent-relative reasons are not shared, it can appear that a deontology defined in terms of such reasons leaves little room for the thought that deontological morality is a shared endeavour -- for instance, if all that matters is that <u>I</u> not perform certain acts, what of my reasons to discourage you from doing them? Second, the focus upon agent-relativity gives rise to a putative worry about the rationale for the legitimate complaints of the victims of wrong-doing -- if a reason is a reason only for a particular agent, how can the victim appeal to this very reason in grounding her complaint against the perpetrator? The third problem which we shall consider is a standard difficulty for the deontologist. Consider, for example, a deontological constraint against killing the innocent. Such a constraint rules out my killing an innocent even if more lives will thereby be saved. From the consequentialist perspective, this manifests unacceptable squeamishness -- I simply want to keep <u>my</u> hands clean. Our purpose here is to adumbrate a deontological framework which meets these three difficulties, and goes some way toward mitigating the apparent "obscurity" (Nagel, p.175) of deontological constraints.

I

<hr/>

[4]. See Nagel's <u>The View from Nowhere</u> and Korsgaard's 'The Reasons We Can Share: An Attack on the Distinction between Agent-Relative and Agent-Neutral Values', <u>Social</u> <u>Philosophy</u> <u>and</u> <u>Policy</u>, **10**, 1993, pp. 24-51. Unless otherwise noted, references to Nagel and Korsgaard are references to these publications.

Before giving a formal account of the distinction between the agent-relative and the agent-neutral it will be useful to outline why this distinction is taken to be crucial in understanding the distinction between deontology and consequentialism. (Since our focus here is to characterise deontology rather than consequentialism, we restrict our remarks, on the whole, to simple direct maximising act-consequentialism). Williams' famous fantasy[1] may serve as an example vis-à-vis constraints. Having been given the choice by the Army Captain of shooting one Indian and thereby saving another nineteen from execution, Jim may decide that it would be wrong for <u>him</u> to shoot the Indian; here his reason is that <u>he</u> ought not to kill an innocent person, even if someone else is going to kill that person if he does not. In the case of options, the suggestion is that it may be permissible for an agent not to take the course of action which would maximise the good, if doing so would have a seriously deleterious effect on the success of <u>his</u> <u>own</u> projects.

The distinction between the agent-relative and agent-neutral is usually, and rightly, seen as purely structural -- it is independent of substantive ontological and epistemological claims. For example, the moral realist might claim truth values for both agent-relative and agent-neutral judgements; and the distinction might equally well be drawn by the projectivist. As rehearsed in the introduction, the distinction is most commonly introduced with respect to reasons: a reason is agent-relative if its full articulation would involve ineliminable pronominal back-reference to the agent; agent-neutral otherwise (see, e.g., Nagel, pp.152-153). The account is often then extended to cover values, without much in the way of an explanation. Korsgaard, for example, assumes (p. 25) an equivalence between values and practical reasons: "to say that there is a practical reason for something is to say that the thing is good and vice versa". This leads her to the, in our view, unfortunate position of claiming that if an agent has a relative reason for promoting some goal then that goal is good for that agent (see, e.g., pp.25-26 and p.47).

We have argued[2] that the definition in terms of reasons, though adequate for many purposes, is not strictly accurate and we prefer a formalisation in terms of rule. As we have discussed the matter elsewhere (see note one) we shall be brief. We treat moral rules as universally quantified directives, governed by a <u>pro tanto</u> operator. Where 'xS' abbreviates, 'x should ensure, to the best of x's abilities,

---

[1]. J.J.C. Smart and B. Williams, <u>Utilitarianism: For and Against</u>, Cambridge, 1973, pp.98-99.

[2]. 'Agent-Relativity and the Doing-Happening Distinction', section II. Our account is not far from Nagel's formal account in <u>The Possibility of Altruism</u>, Oxford University Press, 1970.

and insofar as there is not a conflicting duty of greater weight[3], that', all rules have the following form:

$$(x)(xS [. . .]).$$

A rule of this form is syntactically agent-relative if and only if there is an occurrence of 'x' in the square brackets bound by the <u>initial</u> universal quantifier; agent-neutral otherwise. A rule is <u>genuinely</u> agent-relative if it is syntactically agent-relative, and not equivalent in ethical content to an agent-neutral rule[4] -- rendering the occurrence of 'x' in the square brackets bound by the initial universal quantifier ineliminable. We illustrate the leading idea by three pairs of agent-relative (AR) and agent-neutral (AN) rules, which also serve to introduce another important distinction.

(1) (AR) $(x)(xS [(x$ does not kill the innocent)])

(2) (AN) $(x)(xS [(y)(y$ does not kill the innocent)])

(3) (AR) $(x)(xS [(y)(y$ is a child of x -> y does not starve)])

(4) (AN) $(x)(xS [(y)(z)(y$ is a child of z -> y does not starve)])[5]

(5) (AR) $(x)(xS [(y)(y$ is a child of x -> x loves $y)])$

(6) (AN) $(x)(xS [(y)(z)(y$ is a child of z -> z loves $y)])$

We have suggested[6] that there is a tripartite distinction, orthogonal to the neutral/relative distinction, between what we call author, object, and author-object rules. This distinction applies to both agent-neutral and agent-relative rules, but is most easily grasped in the case of the latter. In the case of agent-relative rules it depends on the way in which reference to the agent enters into the square brackets. (Or, to revert to the account in terms of reasons, on the way in which back-reference

---

[3]. This somewhat lengthy wording is in part a response to a worry raised by Frances Howard-Snyder in correspondence: the fact that you cannot comply perfectly with a rule does not entail that it does not require you to do anything.

[4]. We shall not go into the reasons for the syntactic/genuine agent-relative distinction here: see, e.g., 'Deontology and Agency', section II, for details.

[5]. 'z' is introduced here only to exhibit structural symmetry; strictly speaking the reference is redundant.

[6]. See 'Agent-Relativity and the Doing-Happening Distinction', section III.

to the agent may occur in an agent-relative reason.) (1) is an author agent-relative rule: we have reference, within the square brackets, to the agent as author of his actions; what matters is that <u>he</u> not kill. But there are also rules where what matters is the relation between the agent and some object of moral concern -- (3) is object agent-relative: we have reference, within the square brackets, only to a relation between x and some entity (in this case a child; but this could be any 'entity' of moral concern, such as, for example, a personal project). It is clearly possible to have rules which embody both concerns, as (5) illustrates: here we have a double reference to x, both as author and as involved in a significant relation to an object.

We can now use the relative/neutral distinction to distinguish, in terms of necessary conditions[7], between deontological and direct maximising act-consequentialist theories. A moral theory can be deontological only if it contains at least one genuinely agent-relative rule; it can be direct maximising act-consequentialist only if it contains no agent-relative rules.

Direct maximising act-consequentialism is defined by one over-arching duty: to maximise the good. In the case of a pluralist account of the good, goodness is spelt out in terms of specific <u>pro tanto</u> agent-neutral duties, each of which captures a certain aspect of the good. These duties (e.g., to minimise lying and to minimise the killing of innocents) can conflict; and then they need to be weighed up in order to determine how the overall good is to be maximised. A monist about the good acknowledges only one duty -- e.g., to maximise happiness. On this latter theory, an agent has only to weigh up various courses of action with regard to achieving maximal happiness. On a pluralist view, the agent has to weigh up the various <u>pro tanto</u> agent-neutral duties to find that act which maximises the good.

There are a wide variety of possible deontological systems, but all incorporate agent-relative duties -- and, from the consequentialist perspective, all such duties can require an agent not to maximise the good. Thus, from the consequentialist point of view, they can be seen as constraints on maximising the good. In normal philosophical usage, however, constraints are identified with the narrower class of negative duties (or perfect duties in Kant's sense): duties not to perform certain kinds of action, such as lying, cheating, torturing, killing the innocent, and so on. It is possible for an agent to fulfil each of these duties perfectly (by ensuring that he never (intentionally) acts in violation of them) and it is often held that they cannot come into conflict (because there is always some action I can do which will not fall under any of the relevant descriptions -- an action which will not be a case of lying, or a case of killing the innocent, etc.).

A system can be envisaged which incorporates only negative agent-relative duties, and under which agents are free to pursue their own interests and projects provided that they violate no

---

[7]. Our thanks to Frances Howard-Snyder for emphasising this to us -- see, e.g., her 'The Heart of Consequentialism', <u>Philosophical Studies</u> 76, 107-129, 1994.

constraints. There would be no options, and no need for options, in such a system (options would enter only if the system were thought too stringent, and it was deemed permissible for an agent to violate a constraint if the cost to her of not doing so were too great). However, systems incorporating only negative duties are not ethically plausible: there are surely positive duties in any deontological system, such as (5) above, or an agent-neutral duty such as:

(7)   AN   (x)xS[(y)(y's life is tolerably pleasant)]

But (5) and/or (7) rule out options: they are duties that one has never perfectly fulfilled (in this world, at least), so presumably the trip to Kilimanjaro is ruled out. Whilst we do not view options as a necessary feature of deontology (e.g., W. D. Ross[8] (ch 2) includes a pro tanto duty to maximise the good, thus ensuring there are no options in his system), we are concerned to represent them within our formalisation in such a way as to manifest their agent-relativity.

Options are captured by adding to the system riders on duties, in the manner of the following rider on (7):

(8)   (x)([xE[(y)(y's life is tolerably pleasant)]->x's interests/projects are seriously compromised]
                                        ->[-xS[(y)(y's life is tolerably pleasant)]])

(where 'E' abbreviates 'x ensures, to the best of x's abilities, and insofar as there is not a conflicting duty of greater weight, that').

Whilst (8) is not, by our definition, an agent-relative rule, it is agent-relative in the sense that the underlined occurrence of 'x' is ineliminable. All riders which spell out options will have the same form as (8) (i.e.: (x)([xE[...]->x's interests/projects are seriously compromised]->[-xS[...])), hence options are agent-relative on our account. If there are (unbounded) positive duties, options can be incorporated only by adding a rider to each such duty.

We have identified one type of agent-neutral reason and two types of agent-relative reasons -- that associated with agent-relative duties (type I), and that associated with options (type II). An agent is acting on an agent-neutral reason if she is following an agent-neutral rule. She is acting on a type I agent-relative reason if she is following a genuinely agent-relative rule; and on a type II agent-relative reason if she invokes an agent-relative rider. Of course, simply knowing that an agent is invoking an agent-relative rider tells us nothing of the specific interests and projects she is seeking to protect and/or pursue. But they must be of value to her -- it is here that agent-relative value enters, if anywhere. On Nagel's view, a project might not be of value impersonally construed (if a project has impersonal value, then we have an agent-neutral reason to pursue it), and yet have value-for-the-agent.

It is important to be clear, however, that the contrast between value-for-the-agent and impersonal value does not exhaust the realm of reasons -- it leaves out agent-relative duties (or type I

---

[8]. See, e.g., his The Right and the Good, Oxford, 1930. References to Ross are to this work.

agent-relative reasons). Whilst Nagel explicitly acknowledges this (see, e.g., p.175), he is not always as clear about it as he might be. Consider, for instance:

> If we assign impersonal value to pleasure and pain, then each person can think about his own suffering not just that he has reason to want it gone, but that it's bad and should be got rid of. If on the other hand we limit ourselves to relative reasons, he will have to say that though he has a reason to want an analgesic, there is no reason for him to have one, or for anyone else who happens to be around to give him one. (p.160)

In terms of our way of laying things out, this contrasts the value-for-the-agent of eliminating her pain (if there is a positive duty in the offing, she will have to invoke an agent-relative rider to the effect that she is permitted to alleviate her pain rather than fulfill the duty) with the agent-neutral rule:

(9)   AN   (x)xS[(y)(y is not in pain)].

So Nagel here confines relative reasons to value-for-the-agent (to be fair to him, this is no doubt because he discusses deontological constraints later in the book), omitting the possibility of:

(10)   AR   (x)xS[x does not cause pain directly]

as an agent-relative constraint on (9). (The insertion of 'directly' is to rule out the possibility of an agent's inflicting pain in order to minimise it overall.) (10) gives rise to an agent-relative reason not founded on considerations of agent-relative value.

We raise this point, because it might have misled Korsgaard. For instance, confining relative reasons to value-for-the-agent renders them subjective in at least two senses: first, what is valuable-for-me may well not have impersonal value (indeed, it might detract from impersonal value); second, my reason for valuing what is valuable-for-me might be inaccessible to others, in the sense that they might find that reason unintelligible. This might explain why Korsgaard (p.47) sees agent-relative reasons as "only subjectively normative".

II

Korsgaard focusses on the notion of a shared reason. She argues (see, e.g., pp.24-25 and 47-48) that deontological reasons are not agent-relative, because, she claims, agent-relative reasons are not shareable. In particular, deontology portrays morality as a shared endeavour, so deontological reasons are shared, and hence not agent-relative. Furthermore, if a wrongdoer has only an agent-relative reason for not harming his victim, then the victim has no grounds for complaint, for the wrongdoer's reason is not one shared by the victim and to which she can appeal; it is one which is personal to the agent in a way that precludes others offering it as a reason for him to desist.

We shall contend that Korsgaard mistakenly regards all agent-relative reasons as stemming from agent-relative value. And there is a sense in which agent-relative value cannot be shared which would have pernicious consequences. However, the sense in which type I agent-relative reasons are

not shared has no such consequences. The distinction between the relative and the neutral is not only defensible but, as Korsgaard herself shows, indispensable in understanding deontology[9].

Two or more agents share a reason to perform some particular act if each has a reason to bring about the same state of affairs or achieve the same goal, in virtue of the same features of some situation. Agent-neutral reasons can be shared, or, to put the same point in terms of rules: agent-neutral rules supply agents with a common aim (we take this notion from Parfit -- see note three). If it is bad for children to starve then, we might reasonably claim, everyone has reason to feed children; this is an aim we can all share. While neutral rules give each agent a common aim, relative rules typically give each agent a different aim[10]. My aim, let us suppose, is that my children be educated and your aim is that yours be educated.

Although type I agent-relative reasons typically cannot be shared, object agent-relative reasons are an exception. A mother's reason for ensuring that her daughter is educated, let us suppose, is that she is her child, and that is a reason which she can and does share with her husband, just because they stand in the same relation of parenthood to their daughter. More generally, an object agent-relative reason which applies to some agent in virtue of the relation in which he stands to some object of concern can be shared by anyone who stands in the same relation to the same object of concern.

What of author agent-relative reasons? So long as we think of these as reasons which the agent has as an individual, rather than as a member of a group to which the reason applies, then they cannot be shared at all. Recall:

(1)   (AR)   (x)(xS [(x does not kill the innocent)])

In following (1), Jack has a reason not to kill an innocent person himself, and Jill has a different reason: not to kill an innocent herself.[11] (Jill might have a reason to prevent Jack from killing an

---

[9]. At one point she seems to suggest that she does not really intend to attack the distinction as such. "My quarrel . . . is really with Nagel's account of the source of these reasons, which suggests that values and reasons originate either from personal, idiosyncratic desires or from metaphysical realities of some kind." (p. 26 fn 6) In mounting that attack she does deny, however, that deontological reasons are agent-relative. Since we think this claim is not only false but robs the deontologist of an essential tool in defining his position, her arguments need to be rebutted.

[10]. This is not to say that they do not apply to everyone; the agent-relative rule against killing the innocent is a rule which applies to every agent but which gives each a different aim: that he not kill the innocent even if, thereby, he could prevent more killings of the innocent.

8

innocent, but this is not an author agent-relative reason -- it does not arise from (1).) There is nothing, however, to prevent there being joint author agent-relative reasons (or object-author relative reasons): reasons which we have as a group, rather than as individuals, and these, necessarily, are shared. We slip loosely between talk of quantifying over people and agents, but strictly speaking it is agents of which we are speaking. In this arena, agents are usually individual people, but on occasion they are groups or institutions.

For example, suppose a pair of parents believe it important that each pair of parents read to their own children themselves, but that it does not matter which of them it is; it just has to be a parent. So the parents believe they have a reason for reading to their child themselves which no other person (or group) possesses. Because of the nature of their joint reason, they will have acted in accordance with it provided that one or other or both of them reads to their child. It does not matter which (as it does matter whether Jim or the Army Captain kills the Indian; and as it would matter to each parent if they were both following the straightforward agent-relative rule:

(11)   AR   (x)xS[x reads to x's children]

with quantification over persons). An institutional example might be each company's duty to pay its employees.

The latter pair are examples of author-object agent-relative sharing; examples of strict author agent-relative sharing in the relevant sense are easy to come by in principle, but we can think of no practically plausible examples -- we employ the following merely to illustrate the theoretical point. If in the rule:

(12)   AR   (x)xS[x does not lie]

we quantify over families rather than individuals, then we have an example of author agent-relative sharing. It has the following upshot: if one family member can minimize the number of lies told by the family by lying herself, then she should do so. This demonstrates the impossibility of the universal sharing of author agent-relative duties: if we quantified over a group consisting of all agents, then the latter rule would be equivalent to

(13)   AN   (x)xS[(y)(y does not lie)]

where we quantify over individual members of the group. And this is agent-neutral -- i.e., universal sharing here entails agent-neutrality.

Deontological constraints, as the case of Jim and the Indians illustrates, are typically author agent-relative and not shareable in the Parfitian sense: each agent has a duty not to lie, cheat, steal, kill the innocent etc. herself -- they furnish each of us with a different aim. They are constraints on any policy, even one of achieving the good; it is wrong to act in these ways, even to make the world a

---

[11]. It follows that no two people can share any individual reason which is author-object agent-relative, in virtue of its author agent-relativity.

better place. And, like all agent-relative duties, they apply to everyone and they are requirements; they do not gain their force from the desires, projects or interests of the agent.

Options, on the other hand, do gain their force from the desires, projects and/or interests of the agent. If we admit agent-relative value, they are tied to it. The projects of a particular agent do not (generally speaking) furnish us with a common aim, so they are not shared in the Parfitian sense. Furthermore, they do not furnish us with a common aim because they are tied to value-for-the-agent. This latter point is crucial. Constraints, as we have seen, do not furnish us with a common aim. But, unlike the case with options, it is a mistake to think that this feature of constraints is linked to agent-relative value -- a mistake which can lead one to think that the main concern of the deontologist is to keep his hands clean. For, the fallacious reasoning runs, if agent-relative reasons derive from agent-relative value then constraints must derive from considerations of value-for-the-agent. And the only value-for-the-agent linked to the observation of a constraint is the value-to-her of keeping her hands clean. Thus constraints manifest an unacceptable moral squeamishness. We deny that the source of constraints is agent-relative value.

<center>III</center>

We are now in a position to assess Korsgaard's criticism of the claim that deontological reasons are agent-relative. Her criticisms focus on the example Nagel gives in The View From Nowhere which he uses to try to explain what it is which makes us reluctant to harm the innocent ourselves, even to secure a greater good. Having escaped from a car wreck in which your companions have been seriously injured you reach the sole house in the vicinity, to find that the terrified old lady who lives there has locked herself in her bedroom, which contains the only phone, and refuses to come out. You wonder whether to force her to come out by twisting the arm of her grandchild, hoping that she will emerge in response to his screams. The harm the child will suffer is much less than the harm your companions will suffer if they do not receive medical treatment. Nagel's suggestion is that, were you to twist the child's arm, your action would be immediately guided by evil, even though the act is placed in the wider context of producing a greater good. From your point of view, the harm to your victim looms large. But this is also the case from the victim's point of view; his objection is not simply that a certain amount of harm is done but that he is harmed and used. You have to will to hurt him. Korsgaard does not, apparently, quarrel with the general thrust of Nagel's argument, though she wants to give it a more Kantian twist by suggesting that what is wrong with your action is that you are using someone else as a means to your end. What she denies is that the reason you have can be agent-relative. Her arguments are as follows.

If the reason to which the agent appeals is a relative reason then (1) the "badness of violating it is a badness that is for you" (p.47). But, surely, "when you violate a deontological restriction, it is

<center>10</center>

bad for your victim as well as you." Your victim has a legitimate double objection: to being harmed and to being used. And (2) the victim (and everyone else for that matter) has no grounds for complaint. For a relative reason is a personal reason, a reason only for the person who has it. Others may acknowledge that such a reason exists for the possessor, but they have no reason to ensure that she acts on it, and no grounds for complaint if she doesn't. If the potential wrongdoer's reason is only a reason for her and no-one else, then no-one else, including the victim, has any reason to complain; it is none of the victim's business.

To deal with Korsgaard's second objection first: it apparently rests on a failure to distinguish options from constraints -- that is, Korsgaard mistakenly sees all relative reasons as stemming from agent-relative value. Options are reasons on which the agent is permitted but not required to act. They stem from his own desires and projects which are, in normal cases, his own personal concern.[12] There is thus, quite naturally, normally no particular reason why anyone else should ensure that the agent act on such reasons. It is his own business whether or not he chooses to take up his options. Matters are quite otherwise with constraints. These are <u>requirements</u>, which apply to all and which are not generated by, and may indeed conflict with, the agent's desires and projects. There thus seems every reason why others should be concerned to ensure that the agent acts on such reasons.[13]

What Korsgaard has done, it seems to us, is to confuse two senses in which agent-relative reasons may be personal ("<u>my</u> property", as she puts it on p.47). An agent-relative reason is personal in the Parfitian sense that (normally) only one person can have it; but only reasons generated by options[14], and not those generated by constraints, are personal in the sense that they are solely the business of their proprietor. What Korsgaard wants is a reason for the agent not to twist the child's arm to which others, including the victim, can appeal. But that reason need not be shared. Perhaps she thinks, however, that we can only appeal to reasons we already share. If no-one else can share my

---

[12]. In Section IV of her paper Korsgaard argues that this picture of options, which we find in Nagel, is not adequate. In her discussion of Nagel's account of deontology, however, she employs Nagel's conception of the matter.

[13]. That Korsgaard assimilates constraints to options is shown by the fact that she offers as a comparison to Nagel's example the case where I have a project to climb Kilimanjaro.

[14]. Personal duties to self perhaps also generate reasons that are the business only of their proprietor; we shall not deal with these here.

agent-relative reason, then the victim cannot appeal to it and others can have no reason to care about my acting on it.

This claim is false. First, the agent-relative reason not to twist the child's arm finds expression in the form of an agent-relative rule, on our account; say:

(14)   AR   (x)xS[x does not harm the innocent]

This rule expresses a directive to which everyone, including the victim, is permitted to appeal (this does not speak to any special grounds of complaint that the victim has -- we broach that issue below). Furthermore, others can have (a shared) reason to try to ensure that I act on my reason without their having the same reason as me. Agent-relative reasons (or rules) play a crucial role in deontological ethics, but they are not the only kinds of reasons to which a deontologist can appeal. A deontology which said that each of us has reason not to harm the innocent, but that none of us has any reason whatever to ensure that others don't harm the innocent, would indeed be very unattractive. But this would be artificially to restrict the deontologist's arsenal. For she will typically have neutral as well as relative reasons (or rules) in her system. Among these reasons, on any standard deontological view, will be a shared neutral reason to promote the observance and even enforcement of the rules.

It is not at the first-order level of following moral rules, but at the second-order level, the level at which we educate, encourage and even discipline people into following the first-order rules, that the notion of a shared deontological enterprise in which all have a common aim applies most crucially. The deontologist wants, amongst other things, to minimise the amount of harm to the innocent; but, more importantly, wants to get there by educating people so that they themselves ensure that they refrain. There are moral reasons we can share, but Korsgaard mislocates them.

That there is a shared, agent-neutral, obligation to encourage, through acceptable methods such as education and persuasion, the general observance of the first-order rules, including agent-relative rules, seems highly plausible. (We are not supposing, of course, that everyone should take on the role of ubiquitous moral policeman; this obligation, like every other, is a pro tanto obligation which is sensitive to other considerations. There is plenty of room for tolerance; indeed there may even be a pro tanto obligation to be tolerant in certain circumstances.[15]) We leave the formulation of such agent-neutral rules in our system to the next section.

We are now in a position to deal fairly briefly with Korsgaard's first objection. The particular form in which Korsgaard poses the objection is unfortunate. Even if practical reason always aims at the good (which we deny), it does not follow that if something is a reason for me (but not for others) to act, it can only be because it is good-for-me. Nor, of course, does it follow that if I have an agent-relative reason not to do something this must be because it is bad-for-me. Constraints are agent-

---

[15]. We make this point in response to a criticism levelled by Jonathan Dancy.

relative; yet the advocate of constraints can certainly deny that the violation of a constraint is wrong (solely) because it is bad for the person who does wrong. We take the thrust of her complaint to be, however, that if the reason I have not to twist the child's arm is an agent-relative one, then the victim has nothing to complain about.

We have dealt with part of this point in meeting her second objection. The victim is, like everyone else, permitted to appeal, and even has a duty to promote compliance with, the first order rules. But surely the victim has special grounds for complaint; and our reply has not captured this. The victim clearly does have such special grounds, as is shown by the fact that the wrongdoer owes the victim, but not the rest of us, an apology and reparation[16]. Indeed, if the rest of us complain, it is natural to see us as not just complaining that a constraint has been breached, but complaining on behalf of the victim. What gives the victim this special ground of complaint?

In our view, it has nothing to do with the fact that the rule that has been breached is agent-relative, but rests simply on the fact that the victim has been wronged. Wherever wrongdoing occurs (except in the case of victimless crimes) both the agent and the victim play a special role just in virtue of the fact that one is the wrongdoer and the other has been wronged. And this is the case even if the moral rule that has been breached is an agent-neutral one (e.g., if you are lied to in breach of $(x)xS[(y)(y$ does not lie$)]$, you are wronged and the breacher wrongs you). The victim's complaint is special in that he is being/has been wronged; and it is directed, primarily, to the agent, who should desist, and/or make amends -- this latter primarily to the victim, and not to anyone else.[17] (We say "primarily" because the community does have an interest here and could also have been wronged, but not in the same way as the victim.)

Korsgaard's own positive conclusion is two-fold: first, that there can only be moral reasons within relationships; and second, that these reasons must be shared, and hence cannot be agent-relative.[18] We demur on both counts. To confine moral reasons to relationships denies such reasons to those who are in no morally significant personal relation with the agent or the victim; yet surely the

---

[16]. Thanks to David Owens for pointing us in the right direction here.

[17]. Note that we are not claiming that people can be wronged only if wrong is done -- people can be wronged, and have special grounds for complaint, even in cases where the right act is performed. This explains the regret on the part of an agent even when she knows she did the right thing in breaching a pro tanto rule at some specific person's expense: she does not regret the act per se, but she regrets that she wronged that individual, and might well owe him amends.

[18]. On pp. 48-9 of her article, Korsgaard seems to equate neutral reasons with intersubjective and objective reasons. Since moral reasons are intersubjective they cannot, on her view, be relative.

bystander B can believe that the violation of A's rights by C is wrong, and have a perfectly good moral reason for preventing it, even though A and C stand in no significant moral relation to B (to claim that, e.g., the fact that all three are human generates a morally significant relation, would be to trivialise the claim). And if we abandon the claim that deontological reasons are agent-relative and cannot be shared, then we cannot explain our intuitions in Jim's case -- intuitions which Korsgaard endorses.

<center>IV</center>

We have suggested that, in any sane deontological system, each agent-relative rule prohibiting or enjoining some action will be accompanied by a neutral rule requiring all of us to promote the observance of the rule in question. What form will the neutral rule take and what will the system prescribe if it conflicts with the relative rule? The most obvious suggestion is that to each agent-relative rule there will correspond its standard agent-neutral transform. (To transform an agent-relative rule into its agent-neutral counterpart, simply universally quantify, within the square brackets, over the occurrence(s) of the variable bound by the initial quantifier in the agent-relative rule. (2), (4) and (6) are agent-neutral transforms of (1), (3) and (5) respectively.)

In order for the system to remain deontological in character, there must be some occasions on which, should the relative and neutral versions of the same rule conflict, the relative rule takes precedence. There are at least three ways in which this could be achieved. (a) The rules might be lexically ordered, so that the neutral version of the rule was to be followed only if it did not prevent the agent following the relative version. This way lies rigorism. (b) A weaker ordering principle would give greater weight in general to the relative than to the neutral version of the rule, but would allow that there could be particular circumstances in which, nevertheless, the neutral rule took precedence.[19] (c) The most flexible (and our favoured) position would be to have no general ordering principle at all; the agent has to weigh duties in light of particular circumstances -- with no weighting being ruled out in advance, so that, in particular, the precedence of agent-relative rules is always a possibility; it is this that makes the view deontological.

The suggestion that the neutral rule in question is the standard agent-neutral transform of the corresponding relative rule runs into a difficulty. Take the pair of rules:

(15)   (AR)   (x)(xS [(x does not lie)])

---

[19]. This is the solution Ross canvasses in the case of the principles of promise-keeping and beneficence; the former generally taking precedence over the latter. Similarly for the principles of non-maleficence and beneficence. See Ross, loc cit.

<center>14</center>

(16)   (AN)   (x)(xS [(y)(y does not lie)])

Suppose an agent, A, is tempted to lie to prevent C from telling two lies and that A will do this by obviating the need for C to lie; if A lies then any reason C has to lie lapses. Let us further suppose that A is forbidden to lie by (15) which here takes precedence over (16). A bystander, B, sees that A is tempted to breach (15). Do the rules give B any reason to discourage A from breaching (16)? Quite the contrary, they tell B to encourage A to do so. Rule (15) tells B not to lie <u>himself</u>, but that is not in question here. Rule (16) tells B to minimise the number of lies told. Supposing A is right in his belief that his lying will prevent C from telling two lies, then B is enjoined by rule (16) to encourage A to lie. The result would be a deontological system which both proscribed certain actions and required others to encourage breach of the proscription -- an unhappy position. Is there an additional neutral rule that will do the task?[20] Here is a better candidate:

---

[20]. Note that we are not claiming that (16) has no part to play in a deontological system; only that it will not help us here.

(17)   (AN)   (x)(xS[(y)(yE[y does not lie])])

(17) tells all of us to ensure that each of us ensures that <u>she</u> does not lie. Because it is agent-neutral, (17) gives rise to a reason we can all share.

How does the adoption of (17) get round the difficulty? Might not B reason thus: If I discourage A from telling one lie to prevent two being told by C I shall certainly be acting in accordance with (17) with respect to A. But now how do I apply (17) in respect of C? Doesn't (17) also tell me to ensure that C follows the rule prohibiting lying, and won't I do this if I encourage A and so bring it about that C does not lie? So (17), it might seem, does not tell me to discourage A from lying. This reasoning contains an error. What (17) tells B to do, with respect to C, is to ensure that C ensures that she, C, does not lie. Is B doing that by encouraging A to lie? No. For A's lie will not ensure that <u>C</u> ensures that she does not lie; it ensures, rather, that C has no need to lie. It removes the need for C to ensure that she does not lie. So (17) does tell B not to encourage A to lie, even to prevent more lies being told. Nor does it give A any reason to lie, for A will not thereby be ensuring that C ensures that she does not lie.

Notice that we are adopting a certain reading of (15), by which it instructs each of us not to lie ourselves, but does <u>not</u> entail an obligation (although there is certainly a permission) to cajole others into conformity with it, universal quantification notwithstanding. The obligation to such cajolery enters with (17); so we are claiming, in essence, that (15) does not entail (17). One might dispute this and reason as follows. There is a rule in force here: no one should lie, so each of us should do our best to ensure conformity with it. In some ways, we would welcome an argument to the effect that such a reading is obligatory -- egoism could then be shown incoherent by noting its formulation as:

15

(x)xS[x maximises x's pleasure]

(here 'S' is to be read as the 'should' of rationality in the egoist's sense).

It is no part of egoism that agents should ensure that <u>others'</u> ensure that they maximise <u>their</u> pleasure. However, such a reading would rule out games as well -- seriousness requires:

(x)xS[x wins]

but this does not entail that you should ensure that your opponent ensures that she wins.

Alas, the story does not end with (17). For consider the following (admittedly somewhat unlikely) case: A is intending to lie in order to set an example to C and D, so that each of them will ensure that he (C/D) does not lie (perhaps A is hoping to educate them, by example, as to the <u>pro tanto</u> wrongness of lying). Now, by (17), B is required to encourage A, since she will thereby be maximising the number of agents following (15).

The solution to this difficulty is to note that, just as (15) constrains (16) (on those occasions when (15) takes precedence over (16), agents must not lie in order to minimise the total number of lies told), an agent-relative constraint is required to resist the maximising nature of (17). This is a quite general point: agent-neutral reasons are essentially exhortations to minimise or maximise; deontological resistance to this is always captured structurally by agent-relative rules. In the case of (17), the relevant constraint takes the following form:

(18)   (AR)   (x)xS[x does <u>not</u> <u>directly</u> encourage lying]

(cashing out with precision the notion of direct encouragement is not something we shall attempt; the intent, of course, is to block an interpretation of (18) which leaves B, in not encouraging A, actually in breach of it, because B might be construed as encouraging C and D to lie). (17) and (18), then, are a pair of bystander rules which are fine candidates for inclusion in a deontological system.

To take another example, Arthur Kuflik[1] argues that there can be agent-neutral non-consequentialist theories in which,

> [t]he principles of justice would be treated as basic <u>constraints</u> on conduct which <u>anyone</u> ought to try to get <u>everyone</u> to observe, but which <u>no one</u> is at liberty to violate. A theory of this sort would be agent-neutral, for it would assign the same goal to everyone: that is, 'universal justice by just means only' but it would be non-consequentialist, for the goal would <u>not</u> be to <u>maximize</u> 'the net sum of good minus bad'. (p.798)

However Kuflik is mistaken in supposing such a theory to be agent-neutral. It incorporates the agent-neutral rule:

(19)   AN   (x)xS[(y)yE(y is just)]

But (19) is constrained by:

(20)   AR   (x)xS[x is just].

---

[1]. 'A Defense of Common-Sense Morality', <u>Ethics</u>, **96**, 1986, 784-803

And the combination of (19) constrained by (20) furnishes each of us with a distinct aim: Jack and Jill share the aim of promoting justice, but in doing so Jack must ensure that <u>he</u> is just, and Jill must ensure that <u>she</u> is just. (Note that Kuflik reads constraints such as (20) in the same way that we do -- i.e., he does not see (20) as entailing (19).)

It is our contention that the difference between deontology and consequentialism, which we explicate in terms of the distinction between relative and neutral rules, emerges at every level of moral theory, not just at the first level. All agree that we should, all else equal, encourage people to conform to whatever rules there are (this is one sense in which deontological morality is a shared endeavour). But the deontologist, unlike the consequentialist, thinks that there are agent-relative constraints on the way we can do this.[2]

David McNaughton
Philosophy Dept.
Keele University
Staffs. ST5 5BG
England

Piers Rawling
Philosophy Dept.
University of Missouri-St. Louis
St. Louis
MO 63121-4499
U.S.A

NOTES