



# DIGITAL ACCESS TO SCHOLARSHIP AT HARVARD

## The Reasons We Can Share: An Attack on the Distinction Between Agent-Relative and Agent-Neutral Values

The Harvard community has made this article openly available.  
[Please share](#) how this access benefits you. Your story matters.

<b>Citation</b>	Korsgaard, Christine. 1993. The reasons we can share: An attack on the distinction between agent-relative and agent-neutral values. In <i>Altruism</i> , ed. Ellen Frankel Paul, Fred Dycus Miller, and Jeffrey Paul. Cambridge: Cambridge University Press. Previously published in <i>Social Philosophy and Policy</i> 10, no. 1: 24-51.
<b>Published Version</b>	<a href="http://journals.cambridge.org/action/displayJournal?jid=SOY">http://journals.cambridge.org/action/displayJournal?jid=SOY</a>
<b>Accessed</b>	August 20, 2013 8:28:37 PM EDT
<b>Citable Link</b>	<a href="http://nrs.harvard.edu/urn-3:HUL.InstRepos:3196321">http://nrs.harvard.edu/urn-3:HUL.InstRepos:3196321</a>
<b>Terms of Use</b>	This article was downloaded from Harvard University's DASH repository, and is made available under the terms and conditions applicable to Other Posted Material, as set forth at <a href="http://nrs.harvard.edu/urn-3:HUL.InstRepos:dash.current.terms-of-use#LAA">http://nrs.harvard.edu/urn-3:HUL.InstRepos:dash.current.terms-of-use#LAA</a>

*(Article begins on next page)*

**The Reasons We Can Share**  
An Attack on the Distinction between  
Agent-Relative and Agent-Neutral Values<sup>i</sup>

Christine M. Korsgaard  
Harvard University

Published in *Social Philosophy and Policy*  
Volume 10, Issue 1; January, 1993  
and in *Altruism*,  
edited by Ellen Frankel Paul, Jeffrey Paul, and Fred D. Miller, Jr.  
forthcoming from Cambridge University Press, 1993

To later generations, much of the moral philosophy of the twentieth century will look like a struggle to *escape* from utilitarianism. We seem to succeed in disproving one utilitarian doctrine, only to find ourselves caught in the grip of another. I believe that this is because a basic feature of the consequentialist outlook still pervades and distorts our thinking: the view that the business of morality is to *bring something about*. Too often, the rest of us have pitched our protests as if we were merely objecting to the utilitarian account of *what* the moral agent ought to bring about or *how* he ought to do it. Deontological considerations have been characterized as “side-constraints,” as if they were essentially restrictions on ways to realize ends.<sup>ii</sup> More importantly, moral philosophers have persistently assumed that the primal scene of morality is a scene in which someone does something *to* or *for* someone else. This is the same mistake that children make about another primal scene. The primal scene of morality, I will argue, is not one in which I do something to you or you do something to me, but one in which we do something together. The subject matter of morality is not what we should bring about, but how we should relate to one another. If only Rawls has succeeded in escaping utilitarianism, it is because only Rawls has fully grasped this point. His primal scene, the original position, is one in which a group of people must make a decision together. Their task is to find the reasons they can share.<sup>iii</sup>

In this paper I bring these thoughts to bear on a question which has received attention in recent moral philosophy. In contemporary jargon, the question is whether reasons and values should be understood to be agent-relative or agent-neutral, or whether reasons and values of both kinds exist. In slightly older terms, the question is whether reasons and values are subjective, existing only in relation to individuals; or objective, there for everyone. I begin by explaining the distinction in more detail, and then examine two kinds of examples which have been used to support the claim that values of both kinds must exist. By explicating the structure of the values in these examples, I hope to show that employing the distinction between agent-relative and agent-neutral is not the best way to account for their normative force. Values are neither subjective nor objective, but rather are *intersubjective*. They supervene on the structure of personal relations. <sup>iv</sup>

### I. Agent-Relative and Agent-Neutral Values

In what I have said so far I have assumed an equivalence or at least a direct correspondence 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. In this I follow Thomas Nagel, whose work will be the focus of what I have to say. (VFN 139)<sup>v</sup> Although assuming this equivalence gives us a variety of ways to characterize the distinction in question, it still turns out to be a delicate matter to do so.

According to Nagel, a subjective or agent-relative reason is a reason only for a particular agent to promote something; an objective reason is a reason for anyone to promote the thing.<sup>vi</sup> “Subjective” in this context is not meant to suggest “unreal” or “illusory.” Subjective reasons are real and in one sense universal - they are alike for everyone - but they are personal property. Objective reasons, by contrast, are common property. Formally speaking, a subjective reason exists when the formulation of the reason contains a “free-agent variable” and an objective reason exists when it does

not.(PA 90ff)<sup>vii</sup> Thus suppose we say “There is a reason for any agent to promote *her own* happiness.” This gives me a reason to promote my happiness and you a reason to promote yours, but it does not give you a reason to promote mine or me a reason to promote yours. On the other hand, suppose we say “There is a reason for any agent to promote *any person’s* happiness.” This gives each of us a reason to promote not only her own happiness but the other’s as well.

Formulated in terms of values, it is tempting to say that subjective reasons capture the notion of “Good-For”, while objective reasons capture the notion of “Good-Absolutely”. If there is a reason for any agent to promote *her own* happiness, then my happiness is Good-For me and yours is Good-For you. But if there is a reason for any agent to do what will promote any person’s happiness then any person’s happiness is Good-Absolutely. Human happiness is an objective value which as such makes a claim on all of us. This way of putting the point, however, obscures an important distinction, which I will discuss in the next section.

In *The Possibility of Altruism*, Nagel argued that all subjective reasons and values must be taken to have objective correlates. If it is Good-For me to have something, then we must regard it as Good-Absolutely that I should have it. I cannot do justice to Nagel’s complex argument here, but its central idea can easily be conveyed. Nagel associates a commitment to the objectivity of value with a conception of oneself as one person among others who are equally real. I act on certain considerations which have normative force *for me*: they are subjective reasons. I am capable, however, of viewing myself from an impersonal point of view - as simply a person, one among others who are equally real. When I view myself this way, I still regard these considerations as having normative force.<sup>viii</sup> This is especially clear, Nagel argues, when I consider a situation in which someone else fails to respond to my reasons. This is why we ask “How would you like it if someone did that to you?” when we are trying to get someone to see the normative force of another’s reasons. If I am

tormenting someone, say a stranger, the question invites me to consider the case where a stranger is tormenting me. According to Nagel I should see that I would not merely dislike this, I would also resent it, and my resentment carries with it the thought that my tormentor would have a reason to stop. That reason is the same as my reason for wanting it to stop: that I don't like it. I would expect my tormentor to respond to *my* reason.(PA 82-85) And yet, to a stranger, I am just a person, some person or other. This shows that I view my reasons as having normative force simply insofar as they are a person's reasons, and expect others to do so as well. And that commits me to the view that other people's reasons have normative force for me.<sup>ix</sup> Where there is a subjective reason, then, there is also an objective one, to which everyone should respond.

Later Nagel changed his mind about this conclusion. But before considering that we must ask more exactly what this argument, if it works, establishes.

## II. Two Interpretations of Agent-Neutral Value

Earlier I mentioned that there is a problem with understanding the distinction between relative and neutral values in terms of the distinction between Good-For and Good-Absolutely. The problem is that the claim that something is a reason for everyone may be understood in two different ways, one of which the phrase "Good-Absolutely" tends to conceal.

An agent-neutral value might be a value that is independent of what agents actually value. According to this interpretation, the goodness of, say, my happiness, has what G.E. Moore called an intrinsic value, a property that is independent either of my interest in promoting it or yours.<sup>x</sup> It provides a reason for both of us the way the sun provides light for both of us: because it's out there, shining down. And just as the sun would exist in a world devoid of creatures who see and respond to light, so values would exist in world devoid of creatures who see and respond to reasons. I call this

interpretation of agent-neutral values Objective Realism.<sup>xi</sup> On a less metaphysical view, agent-neutrality does not mean independence of agents as such, but neutrality with respect to the individual identities of agents. On this reading values are intersubjective: they exist for all rational agents, but would not exist in a world without them. I call this view of agent-neutral values Intersubjectivism.<sup>xii</sup>

The difference between these two interpretations of neutral value is naturally associated with two other differences. First, the two views will normally involve a different priority-ordering between subjective or relative and objective or neutral values. According to Objective Realism, subjective values are *derived from* objective ones: an individual comes to value something by perceiving that it has (objective) value. Our relation to values, on this account, is epistemological, a relation of discovery or perception. According to Intersubjectivism, objective values are derived or - better - constructed from subjective ones. Our individual, subjective interests become intersubjective values when, because of the attitude we take towards one another, we come to share each other's ends.<sup>xiii</sup> On this view, our relation to values is one of creation or construction. The second and related difference concerns the possibility of adding and subtracting value across the boundaries between persons. On an Intersubjectivist interpretation, neutral reasons are shared, but they are always initially subjective or agent-relative reasons. So on this view, everything that is good or bad is so because it is good or bad *for* someone. This makes it natural for an Intersubjectivist to deny that values can be added across the boundaries between people. My happiness is good for me and yours is good for you, but the sum of these two values is not good *for* anyone, and so the Intersubjectivist will deny that the sum, as such, is a value.<sup>xiv</sup> But an Objective Realist, who thinks that the value is in the object rather than in its relation to the subject, may think that we can add. Two people's happinesses, both good in themselves, will be better than one. Since consequentialism depends upon the possibility that values may be added, an

Objective realist about value may be a consequentialist, while an Intersubjectivist will not.<sup>xv</sup>

This leaves us with some important questions. We shall want to know how Intersubjectivism could be true, and what there is to choose between it and Objective Realism. These are questions to which I will return in due course. More immediately, I want to raise a question about Nagel. Which kind of agent-neutral values did he intend to defend? This turns out to be a little difficult to establish. In a "Postscript" he later attached to *The Possibility of Altruism*, Nagel says:

This book defends the claim that only objective reasons are acceptable, and that subjective reasons are legitimate only if they can be *derived from* objective ones. I now think that the argument actually establishes a different conclusion: That there are objective reasons corresponding to all subjective ones. It remains possible that the *original subjective reasons from which the others are generated* retain some independent force and are not completely subsumed under them. (PA vii; my emphases)

The first part of this is misleading, since nothing in *The Possibility of Altruism* really requires that subjective reasons be "derived from" objective ones. What the argument establishes (if it works) is that if you are to act in harmony with a conception of yourself as one person among others who are equally real, then you must regard your own and others' subjective reasons and values as being objective as well. This is consistent with the view that the objective values are constructed from - or as Nagel himself says here "generated from" - the subjective ones, and so consistent with an Intersubjectivist interpretation.

In *The View from Nowhere*, Nagel says his project is to bring the method of objectivity to bear on the will.(VFN 4,138) You are first to see to what extent your motives are really reasons, with normative force for you, by seeing to what extent they

may be confirmed or corrected when you view yourself more objectively, as simply a person, one person among others. You are then to see whether these agent-relative reasons can support a still more objective normative force, by considering whether from this point of view they could be taken to have normative force for everyone. This could describe a practical project: the project would be to bring our subjective motives into the impersonal point of view, conferring objective normative force or value upon them as far as that can consistently be done. The result would be Intersubjectivism, and sometimes Nagel sounds as if this is what he has in mind. But at other times he seems to think of it as an epistemological project, one of *discovering* whether what seems to us, subjectively, to be reasons are objectively real. He suggests that we should take reasons to be objectively real if (or to the extent that) the best account of why it seems to us that there are reasons and values is that they are really there. (VFN 141) This sounds a form of Objective Realism, not about Platonic entities of some sort, but about reasons themselves. But it is not perfectly clear what Nagel thinks is involved in the existence of a reason.<sup>xvi</sup> He says that the existence of reasons is dependent on the existence of creatures who can see and respond to reasons:

The reasons are real, they are not just appearances. To be sure, they will be attributed only to a being that has, in addition to desires, a general capacity to develop an objective view of what it should do. Thus, if cockroaches cannot think about what they should do, there is nothing they should do. (VFN 150)

This, however, is in tension with the claims Nagel makes when he is arguing for the existence of neutral values. For instance:

The pain can be detached in thought from the fact that it is mine without losing any of its dreadfulness. It has, so to speak, a life of its own. That is why it is natural to ascribe to it a value of its own. (VFN 160)



... suffering is a bad thing, period, and not just for the sufferer.(VFN 161)

An Intersubjectivist account of neutral values does not require that suffering be a bad thing in itself and not just for the sufferer. It requires only that suffering be a bad thing for everyone *because* it is bad for the sufferer. So here Nagel again seems to be an Objective Realist. But on a Realist conception of the badness of pain, surely the pains of animals who cannot think objectively about what they should do must be bad in the same way as the pain of animals who can. If so, there would be reasons and values, even in a world without creatures who can see and respond to them.<sup>xvii</sup>

Finally, when discussing the temptation to think that a maximally objective account of values must be the best one, Nagel remarks:

This idea underlies the fairly common moral assumption that the only real values are impersonal values, and that someone can really have a reason to do something only if there is an agent-neutral reason for it to happen. That is the essence of traditional forms of consequentialism: the only reason for anyone to do anything is that it would be better in itself, considering the world as a whole, if he did it.(VFN 162-163)

Evidently Nagel thinks that the position that there are only agent-neutral values commits one to consequentialism. Relatedly, he thinks that agent-neutral values are correctly described as reasons for things to happen, reasons that are concerned with what is “better in itself.” This again suggests Objective Realism. Nagel’s position, I think, is not fully consistent. On the whole it seems as if he takes himself defending the existence of agent-neutral reasons in an Objective Realist sense, although his project can be understood as an Intersubjectivist one.

It is not necessary to settle the question of how to categorize Nagel’s position here. But two points are important to the rest of my argument. First, if we distinguish between agent-relative or subjective values on the one hand, and agent-neutral values *understood on the Objective Realist model* on the other, we leave out an

important option. Values may be intersubjective: not part of the fabric of the universe or external truth, but nevertheless shared or at least shareable by agents.<sup>xviii</sup> Second, if the status of values is essentially Intersubjective, then the question arises why we should suppose that a value must be shared by everyone, why Intersubjectivism must be universal. If values arise from human relations, then there are surely more possibilities. The claims springing from an acknowledgement of our common humanity are one source of value, but the claims springing from friendships, marriages, local communities and common interests may be others.

### III. Why Not All Values are Agent-Neutral

By the time he wrote *The View From Nowhere* Nagel had decided that not all subjective values have objective correlates.<sup>xix</sup> He argues that an individual may have agent-relative or subjective reasons which have a legitimate normative force for her but which have no normative force for others.

Nagel was moved to modify his earlier position, I believe, by a general consideration and by reflection on certain familiar categories of value which seem to illustrate that consideration. The general consideration is familiar to us from criticisms of utilitarianism, especially those of Bernard Williams.<sup>xx</sup> According to Williams, utilitarianism deprives the moral agent of her integrity or individual character, because it does not allow her actions to be guided by commitments to a set of people and projects that are her distinctively her own. But these are the very commitments which make us who we are as individuals and give us reasons for caring about our own lives. A person may surely find that some project or person is the most important thing in the world *to her* without having to suppose that it is the most important thing in the world *absolutely*. A theory that requires impartial allegiance to a system of agent-neutral values gives individuals insufficient space in which to lead

their own lives. In Samuel Scheffler's words, it ignores "the independence and distinctness of the personal point of view."<sup>xxi</sup>

In *The View from Nowhere* Nagel discusses three categories of values which, he thinks must be understood as agent-relative for these reasons.(VFN 164ff.) The first category springs from the agent's special relationship to his own projects. Nagel calls these "reasons of autonomy." He gives the example of someone with a desire to climb to the top of Kilimanjaro. This desire, he supposes, could give the person a good reason to make the climb, without giving others a reason to help him to make it.(VFN 167) Because he has the desire, his climbing Kilimanjaro is Good-For him, but this does not make it Good-Absolutely, nor need he suppose that it does. The second category, and the most difficult to understand, is the category of "deontological reasons." These are traditional moral restrictions, which forbid performing certain types of actions even when the consequences of doing so are good. According to Nagel, they spring from an agent's special relationship to his own actions. Although it may be best absolutely that someone should lie or break faith or kill another, because of the good consequences that will in this way be produced, it may be better *for him* not to do so.(VFN 180) The last category is "reasons of obligation" which, Nagel says, "stem from the special obligations we have toward those to whom we are closely related: parents, children, spouses, siblings, fellow members of a community or even a nation."(VFN 165) Because of my special obligation to my own child, for instance, it might be the most important thing in the world *to me* that my child be successful or happy. I can have this attitude without supposing that my child is objectively any more important than any other child.

In each of these three cases, it appears as if an agent has excellent subjective reasons for doing things which from an objective point of view are either completely worthless or obviously inferior to other things which she might do. Of course there are familiar strategies for dealing with these appearances, many of which have been

generated by the utilitarian tradition. The most revisionist is to dismiss them, and castigate people who spend their time on worthless activities as irrational and people who pursue the happiness of their loved ones at the expense of the greater good as selfish. A more moderate strategy is to produce extraneous justifications for giving one's personal concerns extra weight. The good is maximized, say, by everyone looking after her own special friends. But there are also well-known objections to these strategies.<sup>xxii</sup> Rather than supposing that a special concern for your own projects, loved ones, and actions is either irrational or in need of an extraneous justification, Nagel thinks we should allow that there are some values which are purely agent-relative. Accordingly, in *The View from Nowhere*, he offers us explanations of why reasons of the first two kinds, reasons of autonomy and deontological reasons, might be thought to exist. In what follows I examine these accounts.

#### IV. Ambition

In *The View from Nowhere*, Nagel suggests that some of an agent's interests and desires give rise to agent-neutral values and some only to agent-relative values. The obvious question is how we are to draw the line. Nagel expects the two categories to sort along these lines: our interests in avoiding pains and having pleasures, in the satisfaction of what we would intuitively call basic needs, and in the possession of freedom, self-respect, and access to opportunities and resources give rise to neutral values.(VFN 171) But more idiosyncratic personal projects, such as the desire to climb to the top of Kilimanjaro or to learn to play the piano, have only relative value. Rather than using Nagel's label "reasons of autonomy," I am going to call these idiosyncratic projects "ambitions."<sup>xxiii</sup> The claim is that ambitions give those who have them reasons to do things, but do not give others reasons to help or to care whether these things get done. The question then is why the normative force of ambitions is limited in this way.

According to Nagel, it is a matter of how far an individual's authority to confer value may appropriately be thought to extend.(VFN 168) In order to explain this it is helpful to introduce another distinction. Nagel believes that values may differ in what he calls their degree of externality, their independence from the concerns of sentient beings.(VFN 152-153) Some valuable things clearly get their value from their relation to people. Consider for instance chocolate. We could account for the value of chocolate in either of two ways. One is to say that its value is intrinsic, and the reason why we like it so much is because we recognize that fact. If we failed to like chocolate, we would have failed to appreciate something of value. The other is to say that eating chocolate is valuable to human beings *because* we like it so much. In the case of chocolate, that seems like a much more sensible thing to say. Chocolate is not an independent value which our taste buds recognize (as if they were an epistemological faculty, a way of knowing about values). Instead, chocolate gets its value *from* the way it affects us. We *confer value on* it by liking it.

In other cases it is less obvious whether this sort of analysis applies. Consider the value of a beautiful sunset or a work of art. Here, people are much more tempted to say that the value, the beauty, is in the object itself, and that what we do is recognize it. If we didn't like it, we would be failing to see a value that is really there. This is the kind of value that Nagel calls external.<sup>xxiv</sup> Obviously, this kind of value is only possible if we accept an Objective Realist interpretation of agent-neutral or objective values. An Intersubjectivist must say that the value of beauty arises in the same way as value of chocolate, only by a more complex process. In this case, aesthetic value would also be a value that we confer.<sup>xxv</sup>

Leave aside the question whether there are any external values. Suppose that we are talking about those values which we confer. Some of these values are conferred collectively - as aesthetics values are, if they are conferred - while others are conferred individually. This is the phenomenon which Nagel refers to as the

individual's authority. The individual's authority is his right to confer objective value on something by desiring or enjoying or being interested in it. Whenever we say that an agent-neutral value arises from someone's desire, we in effect allow the agent to confer agent-neutral or objective value on some state of affairs. If all desires gave rise to agent-neutral reasons, every desire would be an act of legislation - it would create a value for the whole human race. The question therefore is how far the individual's right to legislate runs: what range of things an individual has the authority to confer neutral or objective value on.

Nagel believes that it is appropriate to give the individual the authority to confer objective value on her own inner states and the conditions that determine what living her life is like, but that it is not appropriate to give an agent the authority to confer objective value on things that are completely outside of herself.(VFN 169-171) Suppose, for example, that it is my ambition that my statue should stand on campus. It seems very odd to say that everyone has a reason to work to bring this about merely because I desire it. Why should I be the right person to determine what state of the campus is objectively good?<sup>xxvi</sup> On the other hand, I seem to be exactly the right person to determine what state of *me* is objectively good. If I'm not the person to determine this, who could possibly be? This is why everyone has a reason to help me to achieve things like pleasure and freedom, but no one has a reason to help me get my statue put up on campus.

Two facts complicate what I have just said, which we must notice in order to avoid confusion. The first is that the satisfaction of a desire often brings pleasure, and Nagel supposes that pleasure has neutral value. So in one sense you do have a reason to help me arrange to get my statue on campus, but it is not, directly, that I want it. It is that, given that I want it, it will give me pleasure. To see that these two reasons are different, we need only remind ourselves that desire and pleasure can be prized apart. We can have desires for the realization of states of affairs in which we will

not personally take part, and desires whose satisfactions we will never even know about.

The other complication comes from one of Nagel's other categories of agent-relative reasons. It seems natural to believe that people have a special obligation to try to promote the projects of those with whom they have personal relationships.(VFN 168) If I am your friend, I should be concerned with whether or not you achieve your ambitions, regardless of whether your doing so serves some objective or neutral value.<sup>xxvii</sup> To correct for these complications, we should imagine a case where all that is relevant is that some randomly selected person has an ambition, and ask whether that ambition, *in itself*, provides others with normative reasons, as it does the person who has it. Suppose I want my statue to stand on campus after I am dead.(VFN 169) I will not be one of those who uses or even sees the campus, nor will I even be around to enjoy the thought that my ambition has been achieved. Someone who takes this desire to be in itself the source of an objectively normative reason must be prepared to let me control campus aesthetics from beyond the grave. According to Nagel, my authority should not extend so far.

This way of putting the question makes Nagel's answer seem reasonable. But it ignores the fact that most people do not regard the value of pursuing their ambitions as grounded merely in their own desires.<sup>xxviii</sup> Here it helps to appeal to a distinction Nagel himself used in *The Possibility of Altruism* - the distinction between unmotivated and motivated desires. (PA 29ff). An unmotivated desire is one which is simply caused in us; a motivated desire is one for which we can give reasons. In *The View from Nowhere*, Nagel says nothing about why his exemplar wants to play the piano or climb to the top of Kilimanjaro.(VFN 167) But most people *do* have reasons for their personal ambitions, and in this sense their ambitions are motivated.<sup>xxix</sup> Attention to this fact reveals that the structure of a reason of ambition is rather complex.

Suppose it is my ambition to write a book about Kant's ethics that will be required reading in all ethics classes. I do not care whether or not I live to see my book required. Following Nagel's analysis we will say that this ambition is agent-relative, since it gives me a reason to try to bring it about that my book is required reading, but it doesn't give anyone else a reason to require my book. This seems to fit, for surely no reason for anyone to require my book could spring from the bare fact that I *want* it that way. The only conceivable reason for anyone to require my book would be that it was a good book.

But this way of describing the situation implies a strange description of my own attitude. It suggests that my desire to have my book required is a product of raw vanity, and that if I want to write a good book, this is merely as a means to getting it required. This does not correctly reflect the structure of my ambition. Part of the reason that I want to write a good book on Kant's ethics is that I think that such a book would be a good thing, and my ambition is not conceivable without that thought. It is an ambition to do something good, and it would not be served by people's requiring my book regardless of whether it was good. For now, let us describe this by saying that I think *someone* should write a book on Kant's ethics good enough that it will be required reading. I think that this would have neutral value.

This doesn't, however, mean that my ambition is just a disinterested response to that neutral value. It is essential not to sanitize the phenomena here, or we shall go wrong. I may be interested in personal adulation, I may really like the idea of my book's being required reading, and I may even harbor competitive feelings towards other engaged in similar projects. I don't just want it to be the case that someone writes the book. I want to *be the someone* who writes that book. That element in my ambition is ineliminably agent-relative; no one else, except possibly my friends, has a reason to care whether I write the book or someone else does.

So the structure of this ambition is not:



i) I want my book to be required reading (where that's an agent-relative end)

ii) therefore: I shall write a good book (as a means to that end)

but rather:

i) Someone should write a book on Kant good enough that it will be required reading. (where that's an agent-neutral end)

ii) I want to be that someone (agent-relative motive)

In other words, to have a personal project or ambition is not to desire a special object which you think is good for you subjectively, but rather to want to stand in a special relationship to something you think is good objectively.

Ambition so characterized clearly does have an agent-relative component: you want to stand in a special relationship to what is good. Is this component the source of subjective normative reasons for action? On the one hand, the agent-relative component does seem to *motivate* me to do a lot of work I would not otherwise do. It is often true that without the personal element in ambition, people would not be able to bring themselves to carry out arduous tasks. There are therefore neutral reasons for encouraging the personal desires associated with ambitions. But should the agent herself treat these personal desires as the sources of reasons? If I took it seriously that my desire that *I* should be the one to write the book was a reason for action, then I would have a reason to prevent one of the other Kant scholars from writing *her* book. But in fact, neither I nor anybody else thinks I have a reason to do this, even if in competitive moments I am tempted to feel it. This is not an expression of ambition, but rather a very familiar perversion of it.

It is important to see that reasons of personal obligation almost always have this form. Although I may not suppose that the happiness of my loved ones is objectively more important than that of anyone else, I certainly do suppose that their happiness is objectively good. The structure of reasons arising from love is similar to that of reasons of ambition. I think that someone should make my darling happy, and

I want very much to *be that someone*. And others may have good reason to encourage me in this. But if I try to prevent someone else from making my darling happy or if I suppose that my darling's happiness has no value unless it is produced by me, that is no longer an expression of love. Again, it is a very familiar perversion of it.<sup>xxx</sup>

Where there is no agent-neutral value anywhere in the structure of the ambition - where the ambition is not an ambition to do something good - we might feel inclined to deny that it provides any kind of a reason, even an agent-relative reason for its agent. This is a plausible way of dealing with my ambition to have my statue on campus. That is just a stupid piece of vanity, and one might well think that such a desire doesn't provide even *me* with a reason for trying to arrange its satisfaction.

But there is an important objection to the way I have handled these cases. I have been trading on the claim that a good book on Kant's ethics would be an objectively good thing. You may of course deny that. But even if you accept it, you might point out that not every ambition is *in that way* an ambition to do or produce something good. Is someone who wants to climb a mountain "because it is there" committed to the view that someone ought to climb this mountain (as if it needed climbing) or perhaps that climbing a mountain an intrinsically valuable action, whose occurrence everyone has a reason to promote? Does someone who wishes to collect stamps or coins or barbed wire, or to excel at bowling or billiards, have to believe that these are activities with an intrinsic value of their own?

Perhaps that does not seem quite right. But neither does it seem right to say that those who pursue such projects are in the grip of unmotivated desires, or view themselves as being so. There are reasons for caring about these things, reasons which are communicable and therefore at least potentially shareable. Ask a mountain climber why she climbs and she need not be mute: she may tell you things about the enlarged vistas, the struggle with the elements, the challenge of overcoming fears or surpassing physical limitations. She takes her desire to climb mountains to be a

motivated desire, motivated by recognizably good features of the experience of climbing. She does not take the value of the climb to be conferred on it simply by her desire to do it. Someone who says “I *just* want to” isn’t offering you his reason; he is setting up a bulwark against incomprehension. You may be the problem or he may feel himself inarticulate: many people do. But listen to the articulate talk about their projects and you hear the familiar voice of humanity, not the voice of alien idiosyncrasies.

Or if you don’t, perhaps you should. For it is at this point that the difference between Objective Realism and Intersubjectivism becomes important. An Objective Realist interpretation of the value of climbing mountains, or of collecting stamps or coins or barbed wire, or of excelling at bowling or billiards, is not very tempting.<sup>xxxii</sup> Neither, as I think, is an Objective Realist interpretation of the value of a good book on Kant’s ethics. These are not intrinsic values, already there in the universe, which we have discovered, but rather are expressions of our own distinctively human capacity to take an interest, and to find something interesting, in whatever we find around us. To share another’s ends, or at least to grant that they could be shared, is to see them as expressions of that capacity, and so as expressions of our common humanity. The Intersubjectivist sees the other as human, and *therefore* shares or tries to share the other’s ends. That is why she helps others to pursue their ambitions. But the Objective Realist sees no reason to help unless he *first* sees the other’s ends as ones that he can share. His relationship to others is mediated by his relationship to their ends. According to the Intersubjectivist this is not only a mistake in moral theory but a moral wrong. We should promote the ends of others not because we recognize the value of those *ends*, but rather out of respect for the humanity of those who have them.

I am not here concerned to argue, as Nagel is in *The Possibility of Altruism*, that we are always obliged to promote everyone’s ambitions, and that therefore we must

find some “combinatorial principles” for weighing up the many reasons they provide. (PA 133ff.) I do not myself believe that reasons *can* be added across the boundaries of persons. And since we cannot always act for everyone’s reasons, that cannot be our duty. But according to this argument we are obliged to see the ends of others as providing reasons for action, and this means that the claims of proximity may bring them into play. Someone in your neighborhood, in immediate need of help in order to carry out his ambition, does present you with a reason to act. In that sense reasons springing from ambitions are agent-neutral. But they spring from our respect for one another, rather than from our respect for one another’s ends.

But one form of proximity is especially important. For of course it is also true that you might come to share the ambition of another in a deeper way. For if what I have said is right, you ought to be committed to the view that another *could* explain to you what is good about the world as she sees it through the eyes of her ambition.<sup>xxxii</sup> You may come to see the value of mountain-climbing, or philosophical ethics, or stamp collecting, and to take it as your own. And then, between the two of you, the value functions *as if* it were a value in the Objective Realist sense. It is a fact about your relationship that you both see this as a good thing, which you share a reason to promote. This is why those who share particular ambitions form communities which acknowledge special and reciprocal obligations to one another. In this way, Intersubjective values can come to function like Objective Realist values with respect to the very communities which they themselves create.

## V. Deontology

Deontological reasons are reasons for an agent to do or avoid certain actions. They do not spring from the consequences of those actions, but rather from the claims of those with whom we interact to be treated by us in certain ways. One who believes in deontological values believes that no matter how good our ends are, we are not

supposed to hurt people, or tell lies, or break promises in their pursuit. Deontological reasons are the source of the traditional moral thou-shalt-nots.

It is important to see why Nagel thinks these reasons must be agent-relative. Three other accounts of them, which construe them as objective or agent-neutral, may seem more plausible at first glance.

First, we might think that they derive directly from the agent-neutral or objectively valuable interests of the other people involved, the potential victims of wrongdoing. We might think that the reason not to hurt people is that it is objectively bad for them to be hurt, or that the reason not to lie to people is that it is objectively good for them to know the truth, or that the reason not to break promises springs from the objective badness of disappointed expectations. In short, we might think that wrong-doing is bad because of the *specific* harm that it does to the victim.

The second account of deontological values is modeled on one utilitarian account of them. John Stuart Mill argued that deontological principles are a kind of inductive generalization from particular utility calculations.<sup>xxxiii</sup> We apply the principle of utility directly in a large number of individual cases, and discover that, almost always, telling lies or breaking promises does more harm than good. Usually, this will be for the kind of reason mentioned in the first account - say, that pain, ignorance, or disappointment is bad - together with certain more long-range considerations, such as the bad effects of setting an example or making a habit of doing such actions.<sup>xxxiv</sup> The actions are bad because of the *general* harm which they do.

Third, we might think that the actions forbidden by deontological reasons are simply bad in themselves, objectively so; not (just) because of the harm that they do but because of a specific form of badness, namely *wrongness*.

But there are problems with all of these attempts to construe deontological values as agent-neutral. To see this consider Bernard Williams's by now famous example:

Jim finds himself in the central square of a small South American town. Tied up against the wall are a row of twenty Indians, most terrified, a few defiant, in front of them several armed men in uniform. ...The captain in charge explains that the Indians are a random group of the inhabitants who, after recent acts of protest against the government, are just about to be killed to remind other possible protestors of the advantages of not protesting. However, since Jim is an honored visitor from another land, the captain is happy to offer him a guest's privilege of killing one of the Indians himself. If Jim accepts, then as a special mark of the occasion, the other Indians will be let off. ...if Jim refuses ...Pedro here will do what he was about to do when Jim arrived, and kill them all.<sup>xxxv</sup>

Utilitarians are committed to the view that it is *obvious* that Jim should kill an Indian, but few people can imagine themselves in Jim's position without some sense of a dilemma. Many think that in Jim's shoes they would kill an Indian, but they do not see it as a happy opportunity for doing some good. Some think that Jim should not let the captain coopt him into participating in a murder and should refuse. Still others think that it is essential to find out, if possible, what the Indians want Jim to do.<sup>xxxvi</sup> Nagel thinks that if all values are objective or agent-neutral we should have *no* sense of dilemma in cases like this, since that we do the *most* good by killing the Indian is obvious.

This problem can be dealt with in various ways. A consequentialist may claim that it is salutary for us to be subject to some hesitation to kill, even when it is irrational. Someone who favors the second account of deontological reasons, in terms of general harm, is especially likely to make this argument: killing is certainly something that usually does more harm than good, so a natural reluctance to do it has a consequentialist value of its own.<sup>xxxvii</sup> Another possible solution is suggested by the

fact that the problem seems to depend on the assumption that values can be added across the boundaries between persons. If we deny this assumption, we may deny that killing twenty Indians is a worse thing than killing one. This move is not open to those who hold that the badness of a wrong act rests in the *general* harm that it does. But those who think that the badness rests in the specific harm to the victim, or in the wrongness of the act itself, may simply refuse to add. According to this view, not only hesitating but refusing to kill the Indian is perfectly intelligible.

But this doesn't entirely solve the problem. Suppose we do think that the badness of killing this Indian rests either in his own resulting death or in the badness of the act of killing him. We refuse to add. Now it looks as if the badness is the same whether the Indian is shot by Jim or by Pedro: there will be a death, and a killing, either way. So perhaps Jim should flip a coin? This doesn't seem right either: most of us think that if Jim doesn't suppose he is going to do any good by killing the Indian then he *certainly* should not kill him. But if the same amount of evil is done either way, then Jim's reason for declining to kill the Indian must be agent-relative.

To make the problem clearer, imagine a peculiar theory of value. According to this theory, value is always objective or agent-neutral, and the only thing that has value is the keeping of promises. This theory will not tell us always to keep our promises, surprising as this may seem. First, assume that we can add values. Then there could be a case like this: by breaking your promise, you could cause five other people to keep theirs; while, if you keep yours, they will break theirs. You produce more promise-keeping by breaking your promise than by keeping it, and so that is what the theory tells you to do. Second, suppose we say that the promise-breakings must be bad for someone, and that their badness cannot be added across the boundaries between persons. For whom are they bad? It doesn't matter which view we take. If the badness is for the victim, I have no reason to care whether I inflict it on him or you

do. I should flip a coin. If the badness is for me, the agent, I may have a reason to care, but it could only be an agent-relative one.

Nagel concludes that deontological reasons, if they exist, are agent-relative. The special relation in which you stand in to an action when you are the one who does it carries a special weight, like the special relations in which you stand in to your own ambitions or loved ones. In taking this position he joins Samuel Scheffler, who had earlier argued that deontological values are agent-relative. In his book, *The Rejection of Consequentialism*, Scheffler argues for what he calls an “agent-centered prerogative,” a right, under certain conditions, to neglect what will conduce to the overall good in favor of one’s personal commitments.<sup>xxxviii</sup> Such a prerogative does the work of Nagel’s “reasons of autonomy.” But Scheffler finds the idea of an “agent-centered restriction” - that is, a deontological requirement - paradoxical. He claims that the idea that there could be a reason not to do certain actions which is not equally a reason to prevent them from being done has “an apparent air of irrationality,” which any account of them must dispel.<sup>xxxix</sup> Although Nagel undertakes to explain how deontological reasons arise, it is clear that he shares Scheffler’s attitude. He characterizes deontological constraints as “obscure” and “peculiar”; he wonders how what we *do* can be so much more important than what *happens*. (VFN 175; 180-181) At one point he says:

One reason for the resistance to deontological constraints is that they are formally puzzling, in a way that the other reasons we have discussed are not. We can understand how autonomous agent-relative reasons might derive from the specific projects and concerns of the agent, and we can understand how neutral reasons might derive from the interests of others, giving each of us reason to take them into account. But how can there be relative reasons to respect the claims of others? How can



there be a reason not to twist someone's arm which is not equally a reason to prevent his arm from being twisted by someone else?

(VFN 178)

Despite his doubts, Nagel gives an account of why they exist.

In cases where a deontological restriction is at issue, doing the *action* puts you into a direct relationship with another human being - your "victim" as Nagel puts it. In performing the action, you will have to aim directly at evil for your victim, even if your larger purpose is good. Nozick, in his remarks on the apparent paradox of deontology, puts the point in more Kantian language. In violating a deontological requirement, you will have to treat your victim as a mere means.<sup>x1</sup> I will come back to the question of what there is to choose between these two formulations. In any case, the force of deontological restrictions, according to Nagel, rests in the immediate badness of victimizing someone.

Nagel illustrates his point with an example.(VFN 176) You need the cooperation of a reluctant elderly woman in order to save someone's life, and you find that you can only secure it by twisting the arm of her grandchild so that his screams will induce her to act. You are faced with using the child as a means to saving a life, and in this case, that involves hurting the child. If the grandmother doesn't give in, you have to try and hurt the child more. You have to *will* to hurt the child more, and so, in a sense to want to.(VFN 182) The louder the child screams, the better for you. But there he is, a child, a vulnerable human being to whom everyone owes protection. From your point of view, this is a terrible thing to *do*.

You might think that this analysis doesn't apply in some of the other cases I've mentioned. Consider Williams's Indians. The one you kill is going to die anyway, whether he is shot by you all alone or along with his compatriots by Pedro. So you are not bringing about an evil for him which he would not have endured otherwise. But there is still a sense in which *you* are aiming directly at his evil. You must pick up a

rifle, aim it at his heart, and fire. You must be gratified if the bullet kills him, just as you must be gratified if the child screams louder. And, despite appearances, there is also a sense in which you are treating him as a mere means. You are killing him in order to save the others. The fact that he is going to die anyway doesn't really change the fact that this is what *you* are doing.

According to Kant, you treat someone as a mere means whenever you treat him in a way to which he could not possibly consent.<sup>xii</sup> Kant's criterion most obviously rules out actions which depend upon force, coercion, or deception for their nature, for it is of the essence of such actions that they make it impossible for their victims to consent. If I am forced I have no chance to consent. If I am deceived I don't know what I am consenting to. If I am coerced my consent itself is forced by means I would reject.<sup>xiii</sup> So if an action depends upon force or deception or coercion it is impossible for me to consent to *it*. To treat someone as an end, by contrast, is to respect his right to use his own reason to determine whether and how he will contribute to what happens.

This is why it is important to establish, if you can, what the Indians themselves think should happen. Suppose the oldest Indian steps forward and says "Please go ahead, shoot *me*, and I forgive you in advance." This doesn't make things wonderful but it does help. *Very* roughly speaking, you are not treating him as a mere means if he consents to what you are doing.<sup>xiiii</sup> Of course the Indian does not in general consent to be shot, and his gesture does not mean that after all he has not been wronged. In the larger moral world he has. But if you and the Indians are forced to regard Pedro and the Captain as mere forces of nature, as in this case you are, then there is a smaller moral world within which the issue is between you and them, and in that world this Indian consents. On the other hand, suppose the Indians are pacifists and they say "We would rather die than ask you, an innocent man, to commit an act of violence. Don't do what the captain asks, but go back up north, and tell our story;

make sure people know what is happening down here.” Now the decision *not* to shoot looks much more tempting, doesn't it? Now you can at least imagine refusing. But you may still take the rifle from Pedro's hands and say “You cannot ask me to kill to save you, and yet I will” and pick an Indian to shoot. This is a different kind of decision to kill than the earlier one, for it involves a refusal to share the Indians' moral universe; from the perspective of the Indians who live, it has a slight taint of paternalism.

Surprisingly, the fact that you are treating someone as a mere means operates even in the peculiar cases of breaking a promise so that other people will keep theirs, or telling a lie so that others will tell the truth. You can see this by imaging the kind of case in which you could be faced with such a decision. If I tell the truth, I predict, three of you will tell lies that you should not tell. On what basis could I make this prediction? Perhaps I think that if I tell the truth I will reveal information which will show you that it is in your interest to lie, and I also think that you are unscrupulous people who will lie if it is in your interest. Or perhaps I believe that the truth will confuse you, and that you will tell the lies as a result of the muddle. Or perhaps I think you have a wrongheaded moral system, and knowing this particular truth will make you wrongly conclude that you ought to lie. However it goes, if I tell a lie in order to *get you* to tell the truth, I am treating you as somehow inferior creatures whose tendency to go wrong must be controlled by my superior wisdom. Since this is a way of being treated to which you could not possibly consent, I am treating you as a mere means. Here I am not necessarily aiming at anything evil for you: I may be paternalistic, protecting you from going wrong. This shows, I think, that Nagel is mistaken when he emphasizes that you are aiming at your victim's *evil*. The problem is that you are treating your victim as a mere means. But suppose that with this revision we accept Nagel's account. It is the particular badness of treating someone as a means that explains deontological reasons. It is the horribleness of looking right into a pair of human eyes, while treating their owner like an piece of furniture or a

tool. And yet by violating the restriction you may be doing what is best. So the badness of violating it is a badness that is for you. The reason is agent-relative.

Now this doesn't seem right at all. Surely when you violate a deontological restriction, it is bad for your victim as well as for you. Your victim may surely object to being treated as a mere means, even when he understands the larger good which is thereby produced. And his objection is not only to being harmed; it is to being *used*. Nagel believes that his theory can accommodate the victim's right to complain. He says:

The deontological constraint permits a victim always to object to those who aim at his harm, and this relation has the same special character of normative magnification when seen from the personal perspective of the victim that it has when seen from the personal perspective of the agent. Such a constraint expresses the direct view of the person on whom he is acting. *It operates through that relation.* The victim feels outrage when he is deliberately harmed even for the greater good of others, not simply because of the quantity of the harm but because of the assault on *his value* of having my actions guided by his evil.(VFN 184; my emphases)

This is absolutely right. But the theory that deontological reasons are agent-relative or only subjectively normative *cannot* accommodate it. If the deontological reason were agent-relative, merely *my* property, my victim would not have the right to demand that I act on it. Consider a comparison. If you have an agent-relative reason to climb Kilimanjaro, and don't do it, I may entertain the thought that you are being irrational. I can see what your reasons are. But if I have no reason to bring it about that you climb Kilimanjaro, as Nagel supposes, then I have no reason to talk you into doing it. I have no reason to do anything about your relative reasons, even to think about them, although I may happen to. I certainly don't have a reason to complain of

your conduct when you don't act on them, and if I do, you may justifiably tell me that it is none of my business. If deontological reasons were *agent*-relative, the same thing would hold for victims. My victim could entertain the thought that I have a reason not to treat him this way, but that thought would give him no grounds for complaint. Astonishingly enough, it turns out to be none of his business.

Earlier, I mentioned two reasons why you might be moved to do something by someone else's subjective or relative reasons. One is to give her the agent-neutral good of pleasure. The other springs from the third category of agent-relative reasons, the reasons of personal obligation. If you stand in a personal relation to someone, you may therefore interest yourself in her subjective reasons. This seems like a natural thing to say, and it has weight against the points I have just been making. Although we may resent it when strangers point out to us that we are not doing what we have reason to do, we do not resent such reminders from friends, and we do not tell them that it is none of their business.

Nagel suggests, in the passage quoted above, that the deontological constraint "operates through the relation" between agent and victim. So it is tempting to suppose that what he has in mind is something like this: The relationship of agents and victims, like that of love or friendship, is a *personal* relationship. Perhaps *that* is what gives the victim a stake in the agent's relative reasons, and so entitles him to complain.<sup>xliv</sup>

But the violation of a deontological constraint *always* involves an agent and a victim, and so if this account is correct, deontological reasons are always shared reasons. They cannot be the personal property of individual agents. Instead, they supervene on the relationships of people who interact with one another. They are intersubjective reasons.

In fact, Nagel's primal scene, the confrontation of agent and victim, shows us how agent-neutral reasons are *created* in personal interaction. My victim complains;

he says: "How would you like it if someone did that to you?" I see not merely that I wouldn't like it, but that I would resent it. I am treating my victim as a means, and it is the essence of treating another as a means that his consent is dispensed with. It would be impossible for me to consent to be so treated and so I would have to rebel. That is why I would feel resentment. "How would you like it if someone did that to you?" In asking me this question my victim demands that I either cease using him as a means, or give up my own claim not to be so used by others. But the latter is impossible: one cannot consent to be used as a means. And so *he obligates me* to desist, and to treat him instead as an end in himself.<sup>xlv</sup> This of course is a variant of Nagel's own argument in *The Possibility of Altruism*.(PA 82ff) And as his arguments there show, my recognition that others must be treated as ends in themselves explains altruistic reasons as well. We resent those who regard our plight with indifference, in much the same way that we resent those who use us as means.

But now we have arrived at a picture of neutral or objective value that is different from the one Nagel had intended to give us. According to this account *all* neutral reasons for action arise from a category which Nagel had thought of as a source of relative reasons - the category of personal relationships. But this is no special category: *forall* human interaction is personal. It is because or to the extent that we regard one another as *persons* that we acknowledge the force of deontological reasons. As persons, others demand that we treat them in ways to which they can consent; as persons, we find we must respond to that demand. But we also express our respect for one another's humanity by sharing in each other's ends. As persons, we have a claim on one another's help when it can readily be given or is desperately needed. It is the status of humanity, as the source of normative claims, that is the source of all value. The argument, in other words, has brought us back to Kant.

## VI. Postscript

Let me conclude by going back to the thoughts with which I began. In both *The Possibility of Altruism* and *The View from Nowhere*, Nagel's arguments take an unexpected turn. In both he starts from recognizably Kantian ideas, working in *The Possibility of Altruism* with motivation derived from a metaphysical conception of the person, and in *The View from Nowhere* with a two-standpoints account. And yet in both he ends up having to construct elaborate arguments to fend off the conclusion that his ideas will lead to utilitarianism. Why does this happen? It happens because Nagel presupposes that the business of morality is to bring something about.<sup>xlvi</sup> This presupposition infects Nagel's arguments in many ways. In *The Possibility of Altruism*, Nagel treats all reasons as reasons to *promote* something.(PA 47ff) In *The View from Nowhere*, he substitutes the idea of aiming at someone's evil for that of treating him as a means. Nagel is puzzled by deontology because he finds it odd that we could have reasons not to do things which are not equally reasons to *prevent* those things from being done.(VFN 177) He does not mention the difference between preventing an action by asking its agent not to do it or talking him out of it and preventing an action by the use of force or tricks. If you suppose that all that matters is what you are bringing about, this is merely a difference in *method*. If morality is concerned with the character of human relationships, this difference is *everything*.<sup>xlvii</sup> It is no accident that in order to explain deontology, Nagel must finally imagine his agents and victims *talking* to each other.<sup>xlviii</sup> Nagel is in danger of ending up with consequentialism because that is where he started.

For the view that the business of ethics is to bring something about is the legacy of utilitarianism, and, in turn, of the scientific aspirations of the utilitarian tradition. According to consequentialist conceptions of ethics, ethics is the most sublime form of technical engineering, the one that tells us how to bring about The Good. The questions that it answers are the questions about what we should do with the world. These *are* the questions we must face when we confront issues of

population control or the preservation of the environment, issues with which utilitarians have been non-accidentally obsessed. But deontological restrictions pre-date these global issues, and were already recognized at a time when all we had to do with the world was to live in it together.

One way in which you might be tempted to describe the position I have defended in this paper leaves the distinction between neutral and relative values in place. It might be thought that I am defending this position: that persons have agent-neutral value; while all other values are agent-relative. And then I add that you express your sense of the neutral value of others by sharing in their agent-relative ends. This *is* close to the Kantian position I want to defend, but it is a misleading way to put it. It makes the value of persons a metaphysical reality, perhaps in need of a metaphysical defense; and to some minds, it will suggest that people are a good thing, and therefore that many people are better than a few. I do not believe these things.

Ask yourself, what is a reason? It is not just a consideration on which you in fact act, but one on which you are supposed to act; it is not just a motive, but rather a normative claim, exerting authority over other people and yourself at other times. To say that you have a reason is to say something *relational*, something which implies the existence of another, at least another self. It announces that you have a claim on that other, or acknowledges her claim on you. For normative claims are not the claims of a metaphysical world of values upon us: they are claims we make on ourselves and each other. It is both the essence of consequentialism and the trouble with it that it treats The Good, rather than people, as the source of normative claims.

The acknowledgement that another is a person is not exactly a reason to treat him in a certain way, but rather something that stands behind the very possibility of reasons. I cannot treat my own impulses to act as *reasons*, rather than mere occurrent impulses, without acknowledging that I at least exist at other times. I cannot treat them as *values*, exerting at least a possible claim on others, without



acknowledging that other persons do indeed exist. That is the lesson of Nagel's own argument in *The Possibility of Altruism*. The title of this paper is a tautology: the only reasons that are possible are the reasons we can share.

---

<sup>i</sup> This paper leaves me with many debts. It is the result of a number of years of teaching Thomas Nagel's books, and I owe a great deal to my students for many helpful comments and pressing challenges. In the fall of 1990, when I was developing my own responses to Nagel in class, I benefited especially from comments by Andrew Livernois and David Sussman; over the course of the last few years, my ideas on these topics have been shaped by conversations with Scott Kim. Arthur Kuflik read drafts of the material at two different stages and commented usefully and extensively both times. I received helpful written comments from many people on an earlier draft, among them James Dreier, Barbara Herman, Andrews Reath, and Amélie Rorty; and benefitted from conversations with Catherine Elgin, Patricia Greenspan, Michael Hardimon, and Samuel Scheffler. Two general discussions - at the Social Philosophy and Policy Center conference for which this paper was written, and with Thomas Scanlon and the members of his seminar on value theory in the Spring of 1992 - provided me with many useful suggestions and clarifications. And I am sure I have been influenced by Stephen Darwall, who makes many of the same points I do in this paper in Part III of his book *Impartial Reason* (Ithaca: Cornell University Press, 1983). But my greatest debt here is of course to Thomas Nagel, whose ideas I have found endlessly fertile even when I have disagreed. I thank all of these people.

---

<sup>ii</sup> The term is used by Robert Nozick in *Anarchy, State, and Utopia* (New York: Basic Books, 1974). I should emphasize that it is the *term* that I am criticizing here. Nozick's account of side-constraints anticipates some of what I will say in this paper about deontological reasons: in particular, that they are based on the Kantian notion that people must not be treated as means (p. 30), and that they will seem puzzling only to someone who assumes that "a moral concern can function only as a moral goal." (p.28)

<sup>iii</sup> John Rawls, *A Theory of Justice* (Cambridge: Harvard University Press, 1971). See especially pp. 139-142.

<sup>iv</sup> This formulation may give rise to the misimpression that I do not think that there can be duties to the self, or that questions of value cannot arise for the self. What I actually think is that the relations between stages of a self have many of the same features as the relations between separate persons; if stages of the self are to lay each other under normative demands, they too owe each other reasons they can share. But, for reasons indicated in Section IV of this paper, it follows that the self cannot have a reason it *could* not in principle share with others. This gives the question of the reasons we can share with others a certain priority, and that is the focus of this paper. Duties to the self do not get an adequate treatment here.

<sup>v</sup> In this paper, references to Nagel's works will be inserted into the text. The abbreviations used are: PA=*The Possibility of Altruism* (Princeton: Princeton University Press, 1970); and VFN = *The View from Nowhere* (New York: Oxford University Press, 1986).

---

<sup>vi</sup> In *The Possibility of Altruism*, Nagel uses the terms “subjective” and “objective.” But these terms are awkward because they are used in so many different ways. “Subjective” may be used in a metaphysical sense, to refer to *how things are for someone*, assuming that things might be different for others. Or it may be used in an epistemological sense, to refer to *how things seem to someone*, assuming that things might in fact be different than they seem. To avoid confusion, notice that in *this* sense the subjective need not be personal or individual. Something could seem the same way to every human being and not be how it is from some more objective point of view. A mirage, although seen by everybody, is in this sense a subjective illusion; more controversially, one might say that colors are a feature of the subjective experience of creatures with color vision. In *The Possibility of Altruism*, Nagel uses “subjective” in a metaphysical sense: a subjective value is one that is Good-For some individual. In *The View from Nowhere*, however, Nagel uses that term to refer to what *seems* to be a reason. Here his project is first to assert that it *seems* to us as if we had reasons and values (from a subjective or personal standpoint), and then raise the question whether, from a more objective or impersonal standpoint, that reveals itself as an illusion of the subjective standpoint or not.(VFN Chapter VIII). For this reason, he borrows Derek Parfit’s terms to cover his earlier distinction. What he had called a subjective value becomes an agent-relative value, which is a source of reasons for a particular agent, but not necessarily for others. What he had called an objective value becomes an agent-neutral value, which is a source of reasons for any agent.(VFN 152)

---

Parfit introduces these terms in *Reasons and Persons* (Oxford: Oxford University Press, 1984), p. 143.

vii James Dreier has pointed out to me that in styling my project an attack on the distinction between relative/subjective and neutral/objective I might give the impression that I think this logical distinction is not exhaustive, which it obviously is. My quarrel, as will emerge, is really with Nagel's account of the source of these reasons, which suggests that values and reasons either originate from personal, idiosyncratic desires or metaphysical realities of some kind. I thank Dreier for the point.

viii More accurately, Nagel's view is that if I do not I will suffer from dissociation between the personal and impersonal views I can take of myself. (PA Chapter XI)

ix Or, as one might put it, that every person, being equally real, is a source of value. But Nagel does not put it that way: he moves, as we shall see, from a focus on the (equal) reality of people to a focus on the reality of their reasons. In one sense, I believe his mistake lies here, and that he would have arrived at a more Kantian and, as I think, more correct position if he had not made this move.

x See especially G.E. Moore, "The Conception of Intrinsic Value" in *Philosophical Studies* (Routledge and Kegan Paul, 1922). Values could still be independent of agents in this sense and still always involve agents in another sense: agents and their experiences might always be parts of the complex "organic unities" which G.E. Moore think are the loci of value. See note 25. I thank Arthur Kuflik for prompting me to be clearer about this.

---

<sup>xi</sup> Another view makes Good-For-ness objective in this sense. It is a fact about the universe that a certain thing is good for me or for you. I *think* that this is the view that G.E. Moore, from whom I borrow the idea of formulating these notions in terms of Good-For and Good-Absolutely, found incoherent. (See *Principia Ethica* (Cambridge: Cambridge University Press, 1903), pp. 97ff). I do not know whether it is incoherent, but it is not tempting.

<sup>xii</sup> For another account of Intersubjectivism see Stephen Darwall, *Impartial Reason* (Ithaca: Cornell University Press, 1983), Part III.

<sup>xiii</sup> It may help to give examples of the sort of position I have in mind here. I am thinking, as will become clear, of Kant's claim that respect for the humanity in the person of another requires you to share his ends; or of Hume's view that the virtues get their value from a shared evaluative standpoint. According to Kant's argument, a person's subjective ends become objective ends in the eyes of those who respect his humanity; according to Hume's, the character traits subjectively valued by the members of a person's own "narrow circle" become objectively valued when viewed from a general point of view which we share. As I suggest below in the text, one may read Nagel's projects as forms of Intersubjectivist constructivism as well. I do not know whether an Intersubjectivist position *must* be one in which objective values are constructed from subjective ones, but the Intersubjectivist positions with which I am familiar do take this form.

<sup>xiv</sup> This is not to say that there cannot be values that are best understood as "good for *us*." But these will not be the results of addition. They will exist when the two of us stand in a relationship to

---

which the value in question is relevant. In this way the birth of a child might be good for a couple, or the conclusion of a treaty might be good for a nation. These are collective, not aggregative, goods.

<sup>xv</sup> Obviously the array of logically possible positions goes far beyond the two that are schematically described in the text. One could be an Intersubjectivist and yet think that values can be added across the boundaries of persons. One could be an Objective Realist and yet deny that values can be added - not only across the boundaries of people, but at all. In *Reasons and Persons* (Oxford: Oxford University Press, 1984), for example, Derek Parfit explores the possibility that weighing and compensation cannot take place even within the boundaries of a life. (pp. 342-345) I am not concerned to discuss all possible theories of neutral value, but only the two I find most natural. I shall assume throughout this paper that if there is any objection to adding values, it comes from the consideration that everything that is good or bad is good or bad for somebody, and that values can be added within individual lives. I shall also assume that the view that everything that is good or bad is good or bad for someone is most naturally associated with some form of intersubjectivism.

<sup>xvi</sup> Sometimes, Nagel seems to imply that all it amounts to for a reason to be “really there” is that it can be assimilated to the objective standpoint without contradiction or incoherence. This unites the practical and the epistemological projects described in the text, and the result would be an Intersubjectivist form of realism. Nagel’s values would be part of reality because we put them there, rather the way that, according to Kant, causes are part of empirical reality. This

---

view would have the merit of giving us realism without metaphysics. But it would require a transcendental argument for the category of objective value, and I do not myself see how, in the absence of Kant's own firm division between theoretical and practical reason, this is to be achieved.

<sup>xvii</sup> Nagel might reply that all that follows is that, if we exist, we have reason to stop the animal's pain. But if pain has a value of its own it seems more natural to say that there just is a reason to stop the animal's pain, although the animal cannot see and respond to it.

<sup>xviii</sup> One reason that I take this option to be important is this: I think that its lack of ontological or metaphysical commitments is a clear advantage of Intersubjectivism; we should not be Objective Realists unless, so to speak, there is no other way. This is not just because of Ockham's razor. A conviction that there are metaphysical truths backing up our claims of value must rest on, and therefore cannot explain, our confidence in our claims of value. Metaphysical moral realism takes us the long way around to end up where we started - at our own deep conviction that our values are not groundless - without giving us what we wanted - some account of the source of that conviction.

<sup>xix</sup> Nagel backed off from his earlier position by degrees. At the time he added the Postscript to *The Possibility of Altruism* (quoted on p. 000), he had decided that it was possible that an individual's subjective reasons may sometimes have a legitimate normative force for her that goes beyond that of their objective correlates. If my happiness is Good Absolutely we both have a reason to pursue it, but perhaps I find an

---

additional or a stronger reason in the fact that it is Good-For me. This seems to be an intermediate position between the views of *The Possibility of Altruism* and *The View from Nowhere*.

<sup>xx</sup> See J.J.C. Smart and Bernard Williams, *Utilitarianism For and Against* (Cambridge: Cambridge University Press, 1973), pp. 100ff.; and Williams, *Moral Luck* (Cambridge: Cambridge University Press, 1981), pp. 1-19.

<sup>xxi</sup> See Samuel Scheffler, *The Rejection of Consequentialism*, (Oxford: Oxford University Press, 1982.) p. 41. I discuss Scheffler's views briefly below, pp. 000-000.

<sup>xxii</sup> Many of which can be found in the two pieces by Williams cited in note 20.

<sup>xxiii</sup> Several readers have pointed out to me that this label, together with the example I go on to discuss, might suggest that all personal projects are in some way competitive. I do not mean to imply that, and in fact discuss some non-competitive ones below. But the choice of an example of a personal project which is competitive seems to me to be useful, since such projects are especially resistant to objectification of either an Objective Realist or an Intersubjectivist kind.

<sup>xxiv</sup> Barbara Herman has pointed out to me that the external account works better for natural beauty than for art, since works of art are socially embedded and therefore their value seems more relative to our interests.

<sup>xxv</sup> One may wonder whether an Objective Realist can accommodate the cases of clearly relational value, like the case of chocolate. The answer is yes. The Objective Realist doesn't have to place the intrinsic value



---

in the chocolate. He can place it in the experience of a human being enjoying eating the chocolate. That is to say, he can construct what G.E. Moore called an organic unity and place the value-creating relationship inside of it. (*Principia Ethica* (Cambridge: Cambridge University Press, 1903), Chapter VI.) The trouble with this strategy is that it conceals the fact that the value is really relational, and the possibility, embraced by Intersubjectivism, that all values are really relational. For further discussion see my “Two Distinctions in Goodness” *Philosophical Review* 92 (1983): 169-195, especially 190-193.

<sup>xxvi</sup> Of course this way of putting it assumes no one else has any desires about the campus that could weigh against mine. In that sense, it assumes that I am the only person in the world who cares about the campus. Some people, when they realize that, are tempted to think that under those improbable circumstances I *would* be the right person to determine what counts as a good state of the campus.

<sup>xxvii</sup> Including your happiness or pleasure, which perhaps makes what I say here controversial. I am claiming that if I care about you I want your ambitions to be fulfilled, and not only in order to make you happy. I want them to be satisfied simply because you do. This is why deathbed wishes are entrusted to loved ones. Of course this does not mean that I will never oppose your pursuit of an ambition if I foresee that it will make you miserable. But that is a matter of weighing, not a matter of refusing to give the ambition any weight of its own. Something here depends on one’s views about rationality. Of course there are people who hold that it is only rational to fulfill those ambitions that will make us happy. If you hold this view about

---

rationality, you are likely to encourage and help your friends only to do what will make them happy, just as you are likely to give up your own more dangerous ambitions. But if you hold that it is sometimes rational just to do what you think is important without regard for your happiness, you are likely to respect a friend's desire to do what he thinks is important without regard to his happiness as well. Of course if you hold the view that happiness just consists in doing what you think is most important, these issues cannot even arise.

<sup>xxviii</sup> On this point see also Stephen Darwall, *Impartial Reason* (Ithaca: Cornell University Press), p. 139.

<sup>xxix</sup> When introducing the idea on PA 29, Nagel writes as if a motivated desire were one arrived at through deliberation. But on his own view prudence is a motivated desire, and most of us can hardly be said to have arrived at through deliberation. You arrive at it through the simple recognition of the reason - that it is your own future - without deliberation. I am using the term in this looser sense; I do not think that most people arrive at their ambitions through deliberation.

<sup>xxx</sup> I am not suggesting that there is something perverted about *sexual* jealousy. The desire to make love to someone is not primarily the desire to be the one who provides him with a certain kind of experience. The desire to make someone happy can be an expression of either morality or of love, but in neither case is it their essence. For further discussion see my "Creating the Kingdom of Ends: Reciprocity and Responsibility in Personal Relations" forthcoming in *Philosophical Perspectives 6: Ethics*, edited by James Tomberlin. Atascadero, California: The Ridgeview Publishing Company, 1992.

---

xxxi One may say that human talents and powers are developed and refined by these activities, and that this is an objective human good. Indeed, when people talk about what they like about these activities, these are the things they will talk about. But this does not mean that what they care about is these supposedly objectively valuable features of their chosen activities rather than the particular activities themselves. Other activities, which the people who value these activities are not always prepared to substitute for them, may refine and develop similar human powers. And one may even accept these other activities as substitutes *if it is necessary* (as when one turns to a less strenuous sport in old age). But we should not take that to mean that the “objective” goods embodied in the activities was all that they cared about. The problem here is like the problem associated with the fact that we love particular people even though what we can say we love about them is general. You love a particular person, not just his warmth, intelligence, and sense of humor. It is not true that any other person with these attributes would do just as well, even though it is true that if he leaves you, you may seek another person with these attributes to replace him. No adequate theory of value can ignore these complex facts.

xxxii There are several ways to motivate this thought. Daniel Warren has pointed out to me, in conversation, that without this thought the requirement to share ends could be met by someone who took a sort of patronizing attitude towards the ambitions of others: “oh, well, *you* like it, so I suppose we shall have to count it as good.” Scott Kim points out that a parallel problem exists on the recipient’s side: if you

---

accept help from someone who does not in any way enter into your ambitions you may be regarding him somewhat instrumentally. The point of these remarks is not to show that there is something *wrong* with either helping or accepting help among those who do not really enter into each other's interests, but that the moral attitude that is required of us is less than perfectly realized. This in turn shows that there is a kind of continuum between the sense of "shared ends" defined in the previous paragraph and the sense defined in this paragraph. One may share the ends of others in the sense of (i) agreeing to promote them because they are another's ends; (ii) trusting that there must really be something interesting about them because they are another's ends; (iii) seeing what is interesting about them; and (iv) coming to have them as your own ends. I thank Thomas Scanlon for prompting me to be clearer about this point, and Amélie Rorty for reminding me of the importance of the possibility that one may stop at step (iii); e.g. one may for instance come to have a much better appreciation of what a certain school of art was trying to do without actually coming to enjoy the works or find them beautiful.

<sup>xxxiii</sup> See John Stuart Mill, *Utilitarianism* (Indianapolis: Hackett Publishing, 1979), pp. 23-24.

<sup>xxxiv</sup> Mill, *Utilitarianism* (Indianapolis: Hackett Publishing, 1979), p. 22.

<sup>xxxv</sup> Williams, in Smart and Williams, *Utilitarianism for and Against* (Cambridge: Cambridge University Press, 1973), p. 98.

<sup>xxxvi</sup> I think this point is sometimes overlooked in discussions of this example. Williams, to be fair, specifies that the Indians are begging Jim to accept the offer. (*Utilitarianism for and Against* (Cambridge:

---

Cambridge University Press, 1973), p. 99.) But he obscures its importance when he says that this is “obviously” what they would be doing.

<sup>xxxvii</sup> This is a familiar move: when reminded that a person is likely to experience a negative moral emotion such as guilt, regret, hesitation, or squeamishness about doing something which according to our theory is right, the philosopher points out that the action in question is *usually* wrong and that it is therefore healthy to be equipped with some reactions which will make it hard for us to do it or will make us think twice before doing it. The assumption seems to be that our emotions are clunkier, more mechanical, less sensitive to the details of a situation, and altogether less refined than our thoughts. This view seems to be a byproduct of the modern conception of the emotions; the emotions are conceived as feelings or reactions, not as perceptions. Aristotle, for instance, would not have said this about the trained emotions of the virtuous person.

<sup>xxxviii</sup> Scheffler, *The Rejection of Consequentialism* (Oxford: Oxford University Press 1982), pp. 14ff.

<sup>xxxix</sup> Ibid. p. 82.

<sup>xl</sup> See note 2.

<sup>xli</sup> See Immanuel Kant, *Grounding for the Metaphysics of Morals* (Hackett, 1983) p.37; Prussian Academy edition p. 430. For interpretation, see my “The Right to Lie: Kant on Dealing with Evil” (*Philosophy and Public Affairs* 15, (1986): pp. 325-349); and Onora O’Neill, “Between Consenting Adults” (*Philosophy and Public Affairs* 14 (1985) : pp. 252-277.

---

<sup>xlii</sup> There are familiar philosophical puzzles about all of these notions. This is perhaps especially true of coercion, notoriously hard to distinguish in any formal way from bribery or the mere offer of an incentive. This is not the place to take these puzzles up, but this should pose a problem only for readers who are actually skeptical about whether there is such a thing as coercion.

<sup>xliii</sup> That is a remark that needs *many* qualifications. Actual consent - in the sense of saying yes - can easily be spurious. As Onora O'Neill argues, a better test of whether someone was able to consent is whether the person had an authentic opportunity to say no. See Onora O'Neill "Justice, Gender, and International Boundaries" forthcoming in *The Quality of Life*, ed. Martha Nussbaum and Amartya Sen. Oxford: The Clarendon Press, 1992.

<sup>xliv</sup> Thomas Scanlon has drawn my attention to a footnote in Nagel's paper "War and Massacre" in which Nagel mentions that Marshall Cohen says that according to Nagel's view, shooting at someone establishes an I-thou relationship. (see *Mortal Questions* (Cambridge: Cambridge University Press, 1979), p. 69). I agree with Cohen - and think that so interpreted Nagel is *right*.

<sup>xlv</sup> Strictly speaking, this is only an account of what Kant would call the "incentive" of morality; we are not obligated until we acknowledge the necessity of adopting this incentive as law. A related point is this: several readers, among them Barbara Herman and Arthur Kuflik, have pointed out to me that this account says nothing about why I must recognize the other as a person, only about what follows from the fact that I do. For now I can only acknowledge that the argument is

---

incomplete in these ways. I hope to say more on these points elsewhere.

<sup>xlvi</sup> A similar point, I think, can be made about Scheffler. He says that it is “natural” to interpret Nozick’s defense of side-constraints as an appeal to the disvalue, the badness, of violating those constraints. (Scheffler, p. 88) But it is only “natural” if you ignore Nozick’s reminder that a moral constraint doesn’t have to function as a moral goal - that is, only if you presuppose that the business of morality is the realization of goals.

<sup>xlvii</sup> Several readers have suggested to me that I am not really rejecting consequentialism but only proposing an alternative account of what we should aim at: decent human relationships. This suggestion is similar to the familiar consequentialist reply to standard counterexamples: “if justice matters, we can include it among the results.” That kind of inclusion results in the curious view discussed on p. 000: that we should commit injustice if it will bring about more justice. Scheffler imagines his consequentialist saying: “And if you are worried that a violation of R [the requirement] corrupts the relationship between the agent and the victim, and that the corruption of a human relationship is a bad thing, then why isn’t it at least as permissible to corrupt one valuable relationship if that is the only way to prevent the corruption of five equally valuable human relationships?”(Scheffler, pp. 89-90) A commitment to mutual respect in human relationships is not merely a commitment to bringing respectful relationships about, any more than a commitment to justice is merely a commitment to bringing justice about. For example: In

---

the early stages of our friendship, I might be tempted conceal things from you in order to help bring about a condition of mutual trust; I might be afraid that you will reject me too quickly if you find certain things out before you know me better. But if mutual trust is ever to be *achieved*, the day must come when my calculations about the effects of my telling you things stops: that is what it *means* for *me* to trust *you*. The point here is that having decent relationships with people is not the same as bringing them about, and to some extent is inconsistent with regarding them as things to be brought about. And my suggestion in this paper is that *having* decent human relationships, not bringing them about, is the primary concern of morality.

<sup>xlviii</sup> Nor is it an accident that many of my own examples in this paper, especially the ones concerning Jim and the Indians, focus on what the protagonists might *say* to each other. Many of Rawls's arguments invite us to imagine people talking to each other, to consider what it would be like to say certain things to another person. His argument against the utilitarian account of what is wrong with slavery, in "Justice as Reciprocity" (in *Utilitarianism with Critical Essays*, ed. Samuel Gorovitz, Bobbs-Merrill, 1971, pp. 242-268) in effect invites us to consider the absurdity of a slaveholder who says to a protesting slave: "But my gains *outweigh* your losses!" His consideration of the effects of publicizing principles of justice on people's self-respect are also related to this theme. (*Theory of Justice*, Harvard University Press, 1971, pp. 177ff.) Part of the appeal of the difference principle is that it is the source of justifications which you can offer *to anyone* without embarrassment.