PHIL 202: Core Ethics  
Fall 2014; Classics in Metaethics  
David O. Brink  
Handout # 8: The Normativity of Ethics III: Nagelian Rationalism

Thomas Nagel's *The Possibility of Altruism* (1970) is one of the few sustained attempts to reject internalist and prudential conceptions of practical reason and to defend the possibility of practical reason that is impartial or altruistic. Nagel makes claims about both moral motivation and practical reason, and each claim has both negative and positive constituents.

- (a) **Motivation**: Nagel appears to reject the need for pro-attitudes in moral motivation (including the Humean view about motivation) and to defend a cognitive account of motivation.
- (b) **Practical Reason**: Nagel both rejects internalist and prudential conceptions of practical reason, including Humean instrumentalism, and defends impartial practical reason.

We have already discussed some of his claims about moral motivation (Handout #2). We will focus here more on his claims about practical reason. Nagel may see these issues as connected if he assumes that practical reason must motivate. If so, he may see (a) as necessary to defend his main claim (b).

1. The rational authority of other-regarding morality requires that practical reason be impartial.
2. Insofar as practical reason is impartial, reasons for action cannot be grounded in an agent’s antecedent desires (or interests).
3. Practical reason motivates in the sense that recognition of reasons for action (normally) motivates.
4. Hence, other-regarding morality is authoritative only if motivation does not require pro-attitudes.

But (1)-(3) do not imply (4) if pro-attitudes can be *consequential* on judgments about practical reason. If so, (b) does not require (a); we can reject the Humean theory of practical reason (instrumentalism) without rejecting the Humean theory of motivation (motivation depends on desire or pro-attitude).

There is also some question about how exactly to understand Nagel’s thesis about altruism. The title of the book signifies a comparatively modest ambition, namely, to establish the *possibility* of altruism. But the structure of the argument of the book (see below) seems to have the more ambitious aim of establishing the *necessity* of altruism -- that is, that other-regarding concern and behavior is a requirement of reason.

In the 1978 postscript, Nagel describes the original aim of the book as a defense of objective reasons against subjective reasons. Nagel clearly regards desire-based reasons as subjective ones. For the most part, I think Nagel treats prudence as a kind of objective reason, and altruism as another kind of objective reason, though there may be places where he treats objective reasons as referring exclusively to altruistic reasons.

- **Objective Monism**: Only objective reasons are defensible; subjective (desire-based) reasons are legitimate only if they are derivable from objective ones.

This involves a sort of eliminativism about subjective reasons.
The thesis which I propose to defend is simply that the only acceptable reasons are objective ones; even if one operates successfully with a subjective principle, one must be able to back it up with an objective principle yielding those same reasons as well as (presumably) others [PA 96; cf. 88, 90, 98].

Looking back, Nagel says he now thinks the argument of the book supports a somewhat more modest conclusion.

- **Objective Dualism**: There are objective reasons corresponding to every subjective reason (and the force of subjective reasons is exhausted by parallel objective reasons).

Even this, he now suspects, is too strong. He now thinks that the force of some subjective reasons may not be exhausted by the corresponding objective ones.

- **Dualism**: There are objective reasons corresponding to every (most?) subjective reasons, but the force of subjective reasons may not be exhausted by parallel objective reasons.

As far as I can see, this leaves the necessity, and not just the possibility, of impartial practical reason intact. What it acknowledges is the possibility that practical reason may not be exclusively impartial and/or objective. He may be leaving room for the possibility of what he calls *reasons of autonomy* in *The View from Nowhere* (1986). For instance, Nagel thinks that my preference to run a marathon gives me (and perhaps those who stand in special relationships to me), but not others, reason to enable me to do so.

**ALTRUISM/IMPARTIALITY**

Nagel takes altruism to be committed to the existence of non-derivative reason to be concerned about and help others.

The general thesis to be defended concerning altruism is that one has *direct* reason to promote the interests of others -- a reason which does not depend on intermediate factors such as one's own interests or one's antecedent sentiments of sympathy and benevolence [PA: 15-16].

Does this mean that altruism requires agent-neutrality? Elsewhere, Nagel distinguishes between agent-relative and agent-neutral reasons.

- **Reasons are agent-relative iff their general form involves essential reference to the agent who has them; otherwise, they are agent-neutral** (*The View from Nowhere*, pp. 152-53).

Agent-neutrality is usually thought to imply a consequentialist normative doctrine in which if one person has a reason to do or care about something (to fear or relieve his own pain) then anyone else has a reason do or care about that thing (e.g. to care about or relieve his pain). By contrast, there are many different possible forms of agent-relativity. Internalist conceptions of practical reason and prudence or egoism are both agent-relative. But so are deontological normative doctrines that recognize side-constraints on promoting the good and C.D. Broad's self-referential altruism, which claims that one has non-derivative reason to benefit anyone it is in one's power to benefit but that the weight of one's reasons to benefit are a function of the relationship between benefactor and beneficiary ("Self and Others").

Nagel himself appears to reject an agent-neutral conception of altruism insofar as he expresses doubts about whether consequentialist normative conceptions can recognize the
separateness of persons and the prohibition on demanding uncompensated sacrifices (PA 138-42).

While some agent-relative conceptions, such as instrumentalism and prudence, deny the existence of nonderivative reason to be concerned about or help others, others, such as self-referential altruism, embrace the existence of nonderivative reason to be concerned about and help others.

Would Nagel regard self-referential altruism as an acceptable conception of altruism? While self-referential altruism does recognize nonderivative reason to be concerned about others, it may not recognize the sort of equal normative concern that Nagel appears to endorse (138-42). His challenge is to find an interpretation of equal concern that does not require the agent-neutrality, which he rejects.

INTERTEMPORAL AND INTERPERSONAL DISTRIBUTION

In contrast with instrumentalism, prudence claims that reasons extend across time; in contrast with prudence, altruism claims that reasons extend across persons.

The interpretation of prudence relates it to the conception of oneself as a temporally persistent being. The failure to accept as reasons for present action the reasons which one expects for the future is the practical expression of a failure to identify with the temporally persistent person of whom one's present self is only a stage among others.

The validation of objective reasons must deal with a transmission of the influence of reasons from one person to another corresponding to the transmission of their influence over time, and this may suggest that if the interpretations are to be parallel, the interpretation of altruism will have to link it to a mystical identification of oneself with other persons, or perhaps with a mass self consisting of all persons. But we have already seen why this is not necessary. To identify with one's future self is not to hold the absurd view that present and future stages of one's life are identical. One need only identify the present as one time among others all of which are contained in a single life. And what corresponds to this in the interpersonal case is not an identification of oneself with other persons or with all persons, but rather a conception of oneself as simply a person among others all of whom are included in a single world [PA 99-100; cf. 19, 58, 89].

Just as the agent's future self provide her with reasons now, so too the interests of others provide her with reasons for actions. Failure to recognize prudence involves temporal dissociation -- failure to recognize the present as one time among others. Failure to recognize altruism involves personal dissociation -- failure to recognize oneself as one among others.

This defense of altruism seems to appeal to the parity of intertemporal and interpersonal distribution -- interpersonal distribution should privilege the agent's own interests no more than intertemporal distribution should privilege the agent's current desires or interests. A proper assessment of this parity argument requires some stage-setting. Sidgwick first raised this parity issue in The Methods of Ethics as a challenge to rational egoism. Nagel appeals to parity to defend the fully neutral view here in The Possibility of Altruism. In Reasons and Persons Parfit appeals to parity to defend a fully relative conception of practical reason that is a form of internalism. Interestingly, Sidgwick thought parity threatened egoism only if we accepted Humean skepticism about personal identity. He thought that the separateness of persons provided a rationale for egoism's hybrid treatment of intertemporal and interpersonal distribution.

To see these issues, we might consider three different structural conceptions of practical reason that vary in their attitudes about the distribution of goods and bade both across persons and across time. A conception is neutral with respect to persons or time if it assigns no intrinsic significance to how goods and bade are distributed among persons or over time. By contrast, we might say, a conception is biased or relative if it does assign intrinsic significance to issues about whom a benefit or burdey falls upon or when it does. Though there could in principle be many
kinds of bias, Parfit focuses on a bias toward the agent himself, in relation to others, and bias toward the present. This would seem to yield four possible conceptions of practical reason.

<table>
<thead>
<tr>
<th>Persons</th>
<th>Time</th>
</tr>
</thead>
<tbody>
<tr>
<td>Neutrality</td>
<td>Neutrality/Rational Benevolence</td>
</tr>
<tr>
<td>Bias</td>
<td>Egoism/Prudence/Self-interest</td>
</tr>
</tbody>
</table>

These are structural conceptions of practical reason that make claims about the distribution of goods and bads across persons and times. They are agnostic about what the good consists in, as between hedonism, desire-satisfaction, perfectionism, or objective list views. Parfit focuses only on three of these four – the fully neutral Rational Benevolence (which he also calls Neutralism), the fully biased or relative Presentism (which he calls the Present-aim Theory), and the hybrid Egoism or Prudence (which he calls the Self-interest Theory). In doing so, he ignores the other hybrid, Benevolence of the Present Moment, perhaps because he thinks it’s a possibility no one has endorsed or wants to endorse.

Part II of Reasons and Persons contains several ingenious arguments against egoism. One of these worries is interestingly related to Nagel’s appeal to parity to argue from instrumentalism to prudence to altruism. This argument appeals to parity against the hybrid structure of egoism. Time and person are parallel distributional dimensions; we need to decide where to locate goods and evils in time and among persons. Once we adopt this perspective, prudence may seem like an unstable hybrid. It says that it makes all the difference on whom a benefit or burden falls and none whatsoever when it falls. On reflection this may seem arbitrary. In The Methods of Ethics Sidgwick considers this issue in the context of his discussion of the proof of utilitarianism.

I do not see why the axiom of Prudence [rational egoism] should not be questioned, when it conflicts with present inclination, on a ground similar to that on which Egoists refuse to admit the axiom of Rational Benevolence. If the Utilitarian [neutralist] has to answer the question, ‘Why should I sacrifice my own happiness for the greater happiness of another?’ it must surely be admissible to ask the Egoist, ‘Why should I sacrifice a present pleasure for a greater one in the future? Why should I concern myself about my own future feelings any more than about the feelings of other persons?’ [418]

The egoist asks the neutralist: Why should I sacrifice my own good for the good of another? The egoist doubts that concern for others is non-derivatively rational. But the presentist can ask the egoist: Why should I sacrifice a present good for myself for the sake of a future good for myself? The presentist doubts that concern for one’s future is non-derivatively rational. These doubts may seem parallel. We must decide where among lives and when within lives to locate goods and harms. Because both are matters of position or location, we may think that they should be treated the same. Parfit pushes this same worry about the hybrid structure of prudence, or the self-interest theory (S), as he calls it, in Part II of Reasons and Persons.

As a hybrid S can be attacked from both directions. And what S claims against one rival may be turned against it by the other. In rejecting Neutralism, a Self-interest Theorist must claim that a reason may have force only for the agent. But the grounds for this claim
support a further claim. If a reason can have force only for the agent, it can have force for the agent only at the time of acting. The Self-interest theorist must reject this claim. He must attack the notion of a time-relative reason. But arguments to show that reasons must be temporally neutral, thus refuting the Present-aim Theory, may also show that reasons must be neutral between different people, thus refuting the Self-interest Theory [140].

If present sacrifice for future benefit is rational, why isn’t sacrifice of one person’s good for the sake of another’s? In this way, the appeal to parity may support neutralism. This is roughly the view Nagel adopts in The Possibility of Altruism. His primary aim is to argue against egoism’s agent-bias and in favor of impartiality or altruism, and he relies on the parity of intertemporal and interpersonal distribution to do so. Just as the interests of an agent’s future self provide him with reasons for action now, so too, Nagel argues, the interests of others can provide him with reason for action. Failure to recognize temporal neutrality involves temporal dissociation -- failure to see the present as just one time among others -- and failure to recognize impartiality or altruism involves personal dissociation -- failure to recognize oneself as just one person among others (16, 19, 99-100).

Alternatively, we might treat time and person as parallel and argue from the agent-bias that egoism concedes to temporal bias, in particular, present-bias. If my sacrifice for another is not rationally required, it may seem that we cannot demand a sacrifice of my current interests for the sake of distant future ones. If so, we will think that it is only the present interests of the agent that provide her with non-derivative reason for action. Though Parfit mentions Nagel’s fully neutral response to parity, it is the fully biased response that he develops and thinks Sidgwick anticipated (RP 137-44).

Whereas Parfit thinks that one cannot defend the hybrid character of prudence, Sidgwick thinks that this challenge to prudence is unanswerable only if we accept Humean skepticism about personal identity over time (Methods 418-19). Sidgwick thinks that prudence is defensible provided we recognize the separateness of persons.

It would be contrary to Common Sense to deny that the distinction between any one individual and any other is real and fundamental, and that consequently "I" am concerned with the quality of my existence as an individual in a sense, fundamentally important, in which I am not concerned with the quality of the existence of other individuals: and this being so, I do not see how it can be proved that this distinction is not to be taken as fundamental in determining the ultimate end of rational action for an individual [498].

This appeal to the separateness of persons suggests a rationale for the hybrid structure of prudence. We saw that when the separateness of persons is invoked to discredit utilitarianism critics of utilitarianism appeal to the compensation principle. But the compensation principle and the metaphysical separateness of persons explain the asymmetry between intrapersonal and interpersonal distribution. We saw that there is automatic intrapersonal compensation but no automatic interpersonal compensation. Compensation requires that beneficiaries also be beneficiaries, and for compensation to be automatic benefactor and beneficiary must be one and the same. In the diachronic, intrapersonal case one’s sacrifice of a present good for a (greater) future good is rational, because there is compensation later for the earlier sacrifice; benefactor and beneficiary are the same. This explains temporal neutrality. But in the interpersonal case, benefactor and beneficiary are different people; unless the beneficiary reciprocates in some way, the agent’s sacrifice will be uncompensated. This explains agent relativity or bias. So we have a rationale for the hybrid treatment prudence accords intertemporal and interpersonal distribution.

Or do we? Couldn’t doubts about interpersonal balancing be extended to intrapersonal balancing? If the separateness of persons defeats interpersonal balancing, why doesn’t the
separateness of different periods within a person’s life defeat intrapersonal balancing? After all, me-now and me-later are distinct parts of me. But then it is hard to see how me-now is any more compensated for its sacrifices on behalf of me-later than I am compensated by my sacrifices for you. Just as doubts about interpersonal balancing lead to a distributed concern with each person, perhaps doubts about intrapersonal balancing should support a distributed concern with each part of a person’s life. There are different interpretations of what this distributed concern requires in the interpersonal context, such as equal distribution and maximin. Perhaps we need to explore comparable interpretations of distributed concern in the intrapersonal context. However, this concern about temporal neutrality is not compelling, as it stands, for several reasons.

First, we might distinguish between temporal impartiality and temporal neutrality. Consider again the interpersonal case. Here, one norm might be called the norm of impartiality; it insists that everyone be given equal concern. This norm of impartiality admits of different interpretations, including a norm of substantive equality and maximin, among others. Indeed, utilitarianism’s person neutrality is one interpretation of interpersonal impartiality. Similarly, we might identify a more generic notion of intertemporal impartiality that would admit of different interpretations, including that of temporal neutrality. One way to read the separateness argument, then, is to see it mandating a temporal impartiality. That would not vindicate temporal neutrality, as such, but it would require a form of impartiality that was inconsistent with the sort of temporal bias displayed in ordinary life by familiar forms of temporal discounting and displayed theoretically in the purebred presentism.

Second, this challenge to temporal neutrality requires thinking that we can and should adopt a sub-personal perspective when reckoning compensation. But there are problems with this idea. Once we go sub-personal and appeal to full relativity, there seems no reason to stop until we reach the sub-personal limit – a momentary time slice of the person. But notions of compensation have no application to momentary time slices, which do not persist long enough to act or receive the benefits of earlier actions. Moreover, many of the goods in life, especially the pursuit and achievement of worthwhile projects, seem to be realized only by temporally extended beings. But if we stop short of momentary time slices and appeal to larger sub-personal entities, call these person segments, other problems arise. One question is just where to stop. If we don’t fully relativize, why relativize partially? Moreover, if we do relativize partially, we introduce indeterminacy. This is because the careers of person segments overlap, with the result that any one point in time is part of the career of indefinitely many different segments. To decide whether compensation has occurred, we need a determinate subject. But if we appeal to person segments, we seem to lack a determinate subject.

Of course, persons are just maximal segments. They also seem to be the most salient segments. Many of the things we value and that structure our pursuits are certain sorts of lives. We aim to be certain sorts of people. Insofar as these ideals structure our beliefs, desires, and intentions, the correct perspective from which to assess success would seem to be the perspective of a whole life. Even when persons have more parochial aims and ambitions, the successful pursuit of these aims and ambitions requires interaction and cooperation among segments, much as persons must often cooperate with others to achieve individual, as well as collective, aims. They do interact and cooperate, much as distinct individuals interact and cooperate in groups, in order to plan and execute long-term projects and goals. They must interact and cooperate if only because they have to share a body and its capacities in order to execute their individual and collective goals, much in the way that individuals must sometimes interact and cooperate if they are to use scarce resources to mutual advantage. Indeed, both the ease and necessity of interaction among person segments will be greater than that among persons, because the physical constraints and the reliability of fellow cooperators are greater in the intrapersonal case. But this means that person segments will overlap with each other; they will stand to each other and the person much as strands of a rope stand to each other and the rope. Though we can recognize the overlapping
strands as entities, the most salient entity is the rope itself. So too, the most salient entity is the person, even if we can recognize the overlapping person segments that make up the person.

In this way, person segments represent a rather arbitrary stopping place. If the appeal to full bias argues for agents with shorter life-spans than persons, then an appeal to full bias ought to argue for person slices as agents. But if, as I have argued, that conception cannot be maintained, then it seems arbitrary to settle on person segments. Once we extend the life-span of the agent beyond that of a person slice, it seems we should keep going until we reach an entity with the most natural borders, viz. the person.

Whereas Parfit defends full relativity, Nagel defends full neutrality, though he is skeptical about the person neutrality of utilitarianism. But intertemporal and interpersonal distribution are importantly disanalogous if, as Nagel appears to believe, there must be compensation for sacrifice (142). Intrapersonal balancing is automatically compensated, because benefactor and beneficiary are the same. But interpersonal balancing is not automatically compensated, because benefactor and beneficiary are distinct.

CONSTANCY OF MOTIVATIONAL CONTENT ACROSS PERSPECTIVES

Nagel believes that motivational content (= authority) must be constant across both temporal and personal perspectives. Consider the prudential case.

Just as the change from a tenseless to a tensed factual judgment does not alter what is believed, but only the standpoint from which one views it, so the change from a tenseless to a tensed practical judgment does not alter what one accepts a justification for wanting, but only the standpoint from which one wants it. If the sense of practical judgments were changed, or their motivational content lost, when one shifted out of the present tense, then practical reasoning would be an area divorced from the conception of oneself as equally real over time [PA 71; cf. 61, 68].

Nagel illustrates the requirement of intertemporal univocity (68-69).

1. I will in six weeks have reason to speak Italian.
2. I now have reason to be able to speak Italian in six weeks.
3. At t6 I have reason to speak Italian.
4. At t6 Brink has reason to speak Italian.

Intertemporal univocity of motivational content is supposed to show that if (1) is true then (2) must be true. Both (1) and (2) are supposed to commit us to the temporally-neutral formulation in (3) and the temporally-neutral and impersonal formulation in (4).

Nagel claims that interpersonal univocity of motivational content involves universality or universalizability.

The avoidance of [interpersonal] dissociation therefore requires the acceptance of universal practical principles which apply in the same sense to everyone, and which are impersonally formulable, so that one can arrive at any true conclusion about what the persons in the situation should do, or have reason to do, without knowing what one’s own place in the situation is, or indeed whether one occupies a place in it at all [108].

But universality doesn’t seem to imply altruism. Altruism’s rival, egoism, also respects universality. Consider this line of reasoning.
1. I have reason to F insofar as it promotes my interest (first-person prudential judgment spoken by me).
2. You have reason to F insofar as it promotes your interest (second-person prudential judgment spoken by my interlocutor).
3. He (David Brink) has reason to F insofar as it promotes his interest (third-person prudential judgment spoken by an observer).
4. Anyone ought to F just insofar as it promotes his own interest (universal egoism).
5. I have a reason to help him F insofar as it promotes his interest (first-person altruistic judgment spoken by me).
6. (x)(y) (If F-ing is in x's interest, then y has reason to help x F) (universal altruism).

(1) expresses the content of the first-person practical judgment (spoken by me, David Brink). This ought to be expressible by a second-person practical judgment (made by someone else), as in (2). And the thought expressed by (1) and (2) ought also to be expressible by a third-person judgment (made by someone not party to the conversation), as in (3). Perhaps these claims commit anyone to a universal third-person practical judgment of the sort expressed in (4). But none of these judgments commit one to the altruistic judgments in (5) or (6). But then it's hard to see how a univocity requirement takes us from judgments that the egoist affirms to altruism.

THE NORMATIVE SIGNIFICANCE OF RESENTMENT

Nagel also makes an intuitive case for altruism on the existence of emotional reactions such as resentment. If you fail to tell me that there is a wasp on my hamburger or refuse to walk around my gouty toes, I am likely to respond to you with resentment (16n, 85). Resentment has normative presuppositions that disappointment, anger, and frustration lack.

- If an earthquake knocks a book off my shelf and onto my gouty toes, I may be pained, frustrated, and angry, but not resentful. There is no one to resent.
- If an earthquake knocks you onto my gouty toes, I may be pained, frustrated, and angry but not resentful.
- If you have Tourette's Syndrome and a tic causes you to tread on my gouty toes, I may be pained, frustrated, and angry, but (probably) not resentful.
- If you knowingly step on my gouty toes, because you couldn't be bothered to veer around them, I will be resentful.

Presumably, what distinguishes resentment from disappointment, anger, and frustration is that it implies, as they do not, that one has been wronged and that someone else has acted as he ought not to have. Moreover, resentment is not predicated on the belief that the offending agent had a self-interested reason not to engage in the destructive behavior. But then it looks like feelings of resentment are evidence against egoism -- the idea that all reasons for action are prudential -- and for altruism.

This argument seems to present a problem of normative accommodation for the egoist that is independent of Nagel's official line of argument that exploits parallels between the justification of prudence and of altruism and that appeals to considerations of universalizability or interpersonal univocity. But Nagel might deny this appearance of independence. He might claim that the egoist who experiences resentment is making judgments that are not interpersonally stable. My feeling of resentment implies the judgment that you (or others) have reason to refrain from causing me unnecessary pain, independently of your own interests. But this requires me to judge that anyone similarly situated (including me) has reason to refrain from causing unnecessary pain to others, independently of his own (my own) interests.
What this shows is that the egoist cannot consistently harbor feelings of resentment toward the inconsiderate. Of course, he has self-interested reason to object to such behavior both prospectively and retrospectively, and he may have self-interested reason to act as if they had necessarily flouted reasons to behave otherwise. But he cannot actually resent them inasmuch as this would imply that they did have and flouted reasons for behaving otherwise. But we can avoid inconsistency by abandoning either egoism or feelings of resentment toward the inconsiderate. Abandoning reactive attitudes such as resentment may seem like a steep price to pay for a philosophical commitment, revealing the unsustainable cost of the egoist living his egoism. Perhaps so. But Nagel also purports to show that egoism is an unstable normative position and that the egoist must recognize the demands of altruism. But the egoist can resist the demands of altruism if he is willing to abandon his attitudes of resentment. The egoist need not be inconsistent, but are we willing to pay the price of being egoists?