Deontological Ethics

First published Wed Nov 21, 2007; substantive revision Wed Dec 11, 2024

The word deontology derives from the Greek words for duty (*deon*) and science (or study) of (*logos*). In contemporary moral philosophy, deontology is one of those kinds of normative theories regarding which choices are morally required, forbidden, or permitted. In other words, deontology falls within the domain of moral theories that guide and assess our choices of what we ought to do (deontic theories), in contrast to those that guide and assess what kind of person we are and should be (aretaic [virtue] theories). And within the domain of moral theories that assess our choices, deontologists—those who subscribe to deontological theories of morality—stand in opposition to *consequentialists*.

---

1. Deontology's Foil: Consequentialism

Because deontological theories are best understood in contrast to consequentialist ones, a brief look at consequentialism and a survey of the problems with it that motivate its deontological opponents, provides a helpful prelude to taking up deontological theories themselves. Consequentialists hold that choices—acts and/or intentions—are not the primary focus of value—that consists of status of affairs. Choices are to be derivatively assessed, in terms of the states of affairs they bring about. Consequentialists thus must specify the states of affairs that are intrinsically valuable—often called, collectively, "the Good." They then are in a position to assert that whatever choices increase the Good, that is, bring about more of it, are the choices that it is morally right to make and to execute. (The Good in that sense is said to be prior to "the Right.")

Consequentialists can and do differ widely in terms of specifying the Good. Some consequentialists are monists about the Good. Utilitarians, for example, identify the Good with pleasure, happiness, desire satisfaction, or "welfare" in some other sense. Other consequentialists are pluralists regarding the Good. Some of such pluralists believe that how the Good is distributed among persons (or all sentient beings) is itself partly constitutive of the Good, whereas conventional utilitarians merely add or average each person's share of the Good to achieve the Good's maximization.

Moreover, there are some consequentialists who hold that the doing or refraining from doing, of certain kinds of acts are themselves intrinsically valuable states of affairs constitutive of the Good. An example of this is

the positing of rights not being violated, or duties being kept, as part of the Good to be maximized—the so-called "utilitarianism of rights" (Nozick 1974).

None of these pluralist positions about the Good erase the difference between consequentialism and deontology. For the essence of consequentialism is still present in such positions as much as it is in monistic positions about the Good: an action would be right only insofar as it maximizes these Good-making states of affairs being caused to exist. Moreover, consequentialists generally agree that the Good is "agent-neutral" (Parfit 1984; Nagel 1986). That is, valuable states of affairs are states of affairs that all agents have reason to achieve without regard to whether such states of affairs are achieved through the exercise of one's own agency or not.

Consequentialism is frequently criticized on a number of grounds. Two of these are particularly apt for revealing the temptations motivating the alternative approach to deontic ethics that is deontology. The two criticisms pertinent here are that consequentialism is, on the one hand, overly demanding, and, on the other hand, that it is not demanding enough. The criticism regarding extreme demandingness runs like this: for consequentialists, there is no realm of moral permissions, no realm of going beyond one's moral duty (supererogation), no realm of moral indifference. All acts are seemingly either required or forbidden. And there also seems to be no space for the consequentialist in which to show partiality to one's own projects or to one's family, friends, and countrymen, leading some critics of consequentialism to deem it a profoundly alienating and perhaps self-effacing moral theory (Williams 1973).

On the other hand, consequentialism is also criticized for what it seemingly permits. It seemingly demands (and thus, of course, permits)

that in certain circumstances innocents be killed, beaten, lied to, or deprived of material goods to produce greater benefits for others. Consequences—and only consequences—can conceivably justify *any* kind of act, for it does not matter how harmful it is to some so long as it is more beneficial to others.

A well-worn example of this over-permissiveness of consequentialism is that of a case standardly called, Transplant. A surgeon has five patients dying of organ failure and one healthy patient whose organs can save the five. In the right circumstances, surgeon will be permitted (and indeed required) by consequentialism to kill the healthy patient to obtain his organs, assuming there are no relevant consequences other than the saving of the five and the death of the one. Likewise, consequentialism will permit (in a case that we shall call, Fat Man) that a fat man be pushed in front of a runaway trolley if his being crushed by the trolley will halt its advance towards five workers trapped on the track. We shall return to these examples later on.

Consequentialists are of course not bereft of replies to these two criticisms. Some retreat from maximizing the Good to "satisficing"—that is, making the achievement of only a certain level of the Good mandatory (Slote 1984). This move opens up some space for personal projects and relationships, as well as a realm of the morally permissible. It is not clear, however, that satisficing is adequately motivated, given that it counterintuitively permits gratuitous preventions of good reasons simply to avoid the problems of maximizing (Bradley 2006). Nor is it clear that the level of mandatory satisficing can be nonarbitrarily specified, or that satisficing will not require deontological constraints to protect satisficers from maximizers.

Another move is to introduce a positive/negative duty distinction within consequentialism. On this view, our (negative) duty is not to make the world worse by actions having bad consequences; lacking is a corresponding (positive) duty to make the world better by actions having good consequences (Bentham 1789 (1948); Quinton 2007). We thus have a consequentialist duty not to kill the one in Transplant or in Fat Man; and there is no counterbalancing duty to save five that overrides this. Yet as with the satisficing move, it is unclear how a consistent consequentialist can motivate this restriction on all-out optimization of the Good.

Yet another idea popular with consequentialists is to move from consequentialism as a theory that directly assesses *acts*, to consequentialism as a theory that directly assesses *rules*—or character-trait inculcation—and assesses acts only indirectly by reference to such rules (or character-traits) (Alexander 1985). Its proponents contend that indirect consequentialism can avoid the criticisms of direct (act) consequentialism because it will not legitimate egregious violations of ordinary moral standards—e.g., the killing of the innocent to bring about some better state of affairs—nor will it be overly demanding and thus alienating each of us from our own projects.

The relevance here of these defensive maneuvers by consequentialists is their common attempt to mimic the intuitively plausible aspects of a non-consequentialist, deontological approach to ethics. For as we shall now explore, the strengths of deontological approaches lie: (1) in their categorical prohibition of actions like the killing of innocents, even when good consequences are in the offing; and (2) in their permission to each of us to pursue our own projects free of any constant demand that we shape those projects so as to make everyone else well off.

2. Deontological Theories

In contrast to consequentialist theories, deontological theories judge the morality of choices by criteria different from the states of affairs those choices bring about. The most familiar forms of deontology, and also the forms presenting the greatest contrast to consequentialism, hold that some choices cannot be justified by their effects—that no matter how morally good their consequences, some choices are morally forbidden. On such familiar deontological accounts of morality, agents cannot make certain wrongful choices even if by doing so the number of those exact kinds of wrongful choices will be minimized (because other agents will be prevented from engaging in similar wrongful choices). For such deontologists, what makes a choice right is its conformity with a moral norm. Such norms are to be simply obeyed by each moral agent; such norm-keepings are not to be maximized by each agent. In this sense, for such deontologists, the Right is said to have priority over the Good. If an act is not in accord with the Right, it may not be undertaken, no matter the Good that it might produce (including even a Good consisting of acts in accordance with the Right).

Analogously, deontologists typically supplement non-consequentialist obligations with non-consequentialist permissions (Scheffler 1982). That is, certain actions can be right even though not maximizing of good consequences, for the rightness of such actions consists in their instantiating certain norms (here, of permission and not of obligation). Such actions are permitted, not just in the weak sense that there is no obligation not to do them, but also in the strong sense that one is permitted to do them even though they are productive of less good consequences than their alternatives (Moore 2008). Such strongly permitted actions include actions one is obligated to do, but (importantly) also included are actions one is not obligated to do. It is this last feature of such actions that warrants their separate mention for deontologists.

2.1 Agent-Centered Deontological Theories

The most traditional mode of taxonomizing deontological theories is to divide them between agent-centered versus victim-centered (or "patient-centered") theories (Scheffler 1988; Kamm 2007). Consider first agent-centered deontological theories. According to agent-centered theories, we each have both permissions and obligations that give us agent-relative reasons for action. An agent-relative reason is an objective reason, just as are agent neutral reasons; neither is to be confused with either the relativistic reasons of a relativist meta-ethics, nor with the subjective reasons that form the nerve of psychological explanations of human action (Nagel 1986). An agent-relative reason is so-called because it is a reason relative to the agent whose reason it is; it need not (although it may) constitute a reason for anyone else. Thus, an agent-relative *obligation* is an obligation for a particular agent to take or refrain from taking some action; and because it is agent-relative, the obligation does not necessarily give anyone else a reason to support that action. Each parent, for example, is commonly thought to have such special obligations to his/her child, obligations not shared by anyone else. Likewise, an agent-relative *permission* is a permission for some agent to do some act even though others may not be permitted to aid that agent in the doing of his permitted action. Each parent, to revert to the same example, is commonly thought to be permitted (at the least) to save his own child even at the cost of not saving two other children to whom he has no special relation. Agent-centered theories and the agent-relative reasons on which they are based not only enjoin each of us to do or not to do certain things; they also instruct me to treat *my* friends, *my* family, *my* promisees in certain ways because they are *mine*, even if by neglecting them I could do more for others' friends, families, and promisees.

At the heart of agent-centered theories (with their agent-relative reasons) is the idea of agency. The moral plausibility of agent-centered theories is

rooted here. The idea is that morality is intensely personal, in the sense that we are each enjoined to keep our own moral house in order. Our categorical obligations are not to focus on how our actions cause or enable other agents to do evil; the focus of our categorical obligations is to keep our own agency free of moral taint.

Each agent's distinctive moral concern with his/her own agency puts some pressure on agent-centered theories to clarify how and when our agency is or is not involved in various situations. Agent-centered theories famously divide between those that emphasize the role of intention or other mental states in constituting the morally important kind of agency, and those that emphasize the actions of agents as playing such a role (Wedgewood 2011). There are also agent-centered theories that emphasize both intentions and actions equally in constituting the morally relevant agency of persons.

On the first of these three agent-relative views, it is most commonly asserted that it is our intended ends and intended means that most crucially define our agency. Such intentions mark out what it is we set out to achieve through our actions. If we intend something bad as an end, or even as a means to some more beneficent end, we are said to have "set ourselves at evil," something we are categorically forbidden to do (Aquinas *Summa Theologica*; Nagel 1979).

Three items usefully contrasted with such intentions are *belief*, *risk*, and *cause*. If we predict that an act of ours will result in evil, such prediction is a cognitive state (of belief); it is not a conative state of intention to bring about such a result, either as an end in itself or as a means to some other end. In this case, our agency is involved only to the extent that we have shown ourselves as being willing to tolerate evil results flowing from our acts; but we have not set out to achieve such evil by our acts. Likewise, a risking and/or causing of some evil result is distinct from any intention to achieve it. We can intend such a result, and we can even execute such an

intention so that it becomes a trying, without in fact either causing or even risking it. (It is, however, true that we must *believe* we are risking the result to some extent, however minimal, for the result to be what we intend to bring about by our act.) Also, we can cause or risk such results without intending them. For example, we can intend to kill and even try to kill someone without killing him; and we can kill him without intending or trying to kill him, as when we kill accidentally. Intending thus does not collapse into risking, causing, or predicting; and on the version of agent-centered deontology here considered, it is intending (or perhaps trying) alone that marks the involvement of our agency in a way so as to bring agent-centered obligations and permissions into play.

Deontologists of this stripe are committed to something like the doctrine of double effect, a long-established doctrine of Catholic theology (Woodward 2001). The Doctrine in its most familiar form asserts that we are categorically forbidden to intend evils such as killing the innocent or torturing others, even though doing such acts would minimize the doing of like acts by others (or even ourselves) in the future. By contrast, if we only risk, cause, or predict that our acts will have consequences making them acts of killing or of torture, then we might be able to justify the doing of such acts by the killing/torture-minimizing consequences of such actions. Whether such distinctions are plausible is standardly taken to measure the plausibility of an intention-focused version of the agent-centered version of deontology.

There are other versions of mental-state focused agent relativity that do not focus on intentions (Hurd 1994). Some of these versions focus on predictive belief as much as on intention (at least when the belief is of a high degree of certainty). Other versions focus on intended ends ("motives") alone. Still others focus on the deliberative processes that precede the formation of intentions, so that even to contemplate the doing of an evil act impermissibly invokes our agency (Anscombe 1958; Geach

1969; Nagel 1979). But intention-focused versions are the most familiar versions of so-called "inner wickedness" versions of agent-centered deontology.

The second kind of agent-centered deontology is one focused on actions, not mental states. Such a view can concede that all human actions must originate with some kind of mental state, often styled a volition or a willing; such a view can even concede that volitions or willings are an intention of a certain kind (Moore 1993, Ch. 6). Indeed, such source of human actions in willing is what plausibly connects actions to the agency that is of moral concern on the agent-centered version of deontology. Yet to will the movement of a finger on a trigger is distinct from an intention to kill a person by that finger movement. The act view of agency is thus distinct from the intentions (or other mental state) view of agency.

On this view, our agent-relative obligations and permissions have as their content certain kinds of actions: we are obligated not to *kill* innocents for example. The killing of an innocent of course requires that there be a death of such innocent, but there is no agency involved in mere events such as deaths. Needed for there to be a *killing* are two other items. One we remarked on before: the action of the putative agent must have its source in a willing. But the other maker of agency here is more interesting for present purposes: the willing must *cause* the death of the innocent for an act to be a *killing* of such innocent. Much (on this view) is loaded into the requirement of causation.

First, *causings* of evils like deaths of innocents are commonly distinguished from *omissions* to prevent such deaths. Holding a baby's head under water until it drowns is a killing; seeing a baby lying face down in a puddle and doing nothing to save it when one could do so easily is a failure to prevent its death. Our categorical obligations are usually negative in content: we are not to kill the baby. We may have an obligation to save it, but this will not be an agent-relative obligation, on the view here considered, unless we have some special relationship to the baby.

Second, *causings* are distinguished from *allowings* (Woollard, 2015, 2023). In a narrow sense of the word we will here stipulate, one *allows* a death to occur when: (1) one's action merely removes a defense the victim otherwise would have had against death; and (2) such removal returns the victim to some morally appropriate baseline (Kamm 1994, 1996; MacMahan 2003). Thus, mercy-killings, or euthanasia, are outside of our deontological obligations (and thus eligible for justification by good consequences) so long as one's act: (1) only removes a defense against death that the agent herself had earlier provided, such as disconnecting medical equipment that is keeping the patient alive when that disconnecting is done by the medical personnel that attached the patient to the equipment originally; and (2) the equipment could justifiably have been hooked up to another patient, where it could do some good, had the doctors known at the time of connection what they know at the time of disconnection.

Third, one is said not to *cause* an evil such as a death when one's acts merely *enable* (or aid) some other agent to cause such evil (Hart and Honore 1985). Thus, one is not categorically forbidden to drive the terrorists to where they can kill the policeman (if the alternative is death of one's family), even though one would be categorically forbidden to kill the policeman oneself (even where the alternative is death of one's family) (Moore 2008). Nor is one categorically forbidden to select which of a group of villagers shall be unjustly executed by another who is pursuing his own purposes (Williams 1973).

Fourth, one is said not to *cause* an evil such as a death when one merely redirects a presently existing threat to many so that it now threatens only one (or a few) (Thomson 1985). In the time-honored example of the run-

away trolley (Trolley), one may turn a trolley so that it runs over one trapped workman so as to save five workmen trapped on the other track, even though it is not permissible for an agent to have initiated the movement of the trolley towards the one to save five (Foot 1967; Thomson 1985).

Fifth, our agency is said not to be involved in mere *accelerations* of evils about to happen anyway, as opposed to *causing* such evils by doing acts necessary for such evils to occur (G. Williams 1961; Brody 1996). Thus, when a victim is about to fall to his death anyway, dragging a rescuer with him too, the rescuer may cut the rope connecting them. Rescuer is accelerating, but not causing, the death that was about to occur anyway.

All of these last five distinctions have been suggested to be part and parcel of another centuries-old Catholic doctrine, that of the doctrine of doing and allowing (see the entry on doing vs. allowing harm) (Moore 2008; Kamm 1994; Foot 1967; Quinn 1989). According to this doctrine, one may not *cause* death, for that would be a killing, a "doing;" but one may fail to prevent death, allow (in the narrow sense) death to occur, enable another to cause death, redirect a life-threatening item from many to one, or accelerate a death about to happen anyway, if good enough consequences are in the offing. As with the Doctrine of Double Effect, how plausible one finds these applications of the doctrine of doing and allowing will determine how plausible one finds this cause-based view of human agency.

A third kind of agent-centered deontology can be obtained by simply conjoining the other two agent-centered views (Hurd 1994). This view would be that agency in the relevant sense requires both intending and causing (i.e., acting) (Moore 2008). On this view, our agent-relative obligations do not focus on causings or intentions separately; rather, the content of such obligations is focused on *intended causings*. For example,

our deontological obligation with respect to human life is neither an obligation not to kill nor an obligation not to intend to kill; rather, it is an obligation not to *murder*, that is, to kill in execution of an intention to kill.

By requiring both intention and causings to constitute human agency, this third view avoids the seeming overbreadth of our obligations if either intention or action alone marked such agency. Suppose our agent–relative obligation were not to do some *action* such as kill an innocent—is that obligation breached by a merely negligent killing, so that we deserve the serious blame of having breached such a categorical norm (Hurd 1994)? (Of course, one might be somewhat blameworthy on consequentialist grounds (Hurd 1995), or perhaps not blameworthy at all (Moore and Hurd 2011).) Alternatively, suppose our agent-relative obligation were not to intend to kill—does that mean we could not justify forming such an intention when good consequences would be the result, *and* when we are sure we cannot act so as to fulfill such intention (Hurd 1994)? If our agent-relative obligation is neither of these alone, but is rather, that we are not to kill in execution of an intention to kill, both such instances of seeming overbreadth in the reach of our obligations, are avoided.

Whichever of these three agent-centered theories one finds most plausible, they each suffer from some common problems. A fundamental worry is the moral unattractiveness of the focus on self that is the nerve of any agent-centered deontology. The importance of each person's agency to himself/herself has a narcissistic flavor to it that seems unattractive to many. It seemingly justifies each of us keeping our own moral house in order even at the expense of the world becoming much worse. The worry is not that agent-centered deontology is just another form of egoism, according to which the content of one's duties exclusively concern oneself; even so, the character of agent-relative duties is such that they betoken an emphasis on self that is unattractive in the same way that such emphasis makes egoism unattractive. Secondly, many find the distinctions

invited by the Doctrine of Double Effect and the (five versions of the) Doctrine of Doing and Allowing to be either morally unattractive or conceptually incoherent. Such critics find the differences between intending/foreseeing, causing/omitting, causing/allowing, causing/enabling, causing/redirecting, causing/accelerating to be morally insignificant. (On act/omission (Rachels 1975); on doing/allowing (Kagan 1989); on intending/foreseeing (Bennett 1981; Davis 1984).) They urge, for example, that failing to prevent a death one could easily prevent is as blameworthy as causing a death, so that a morality that radically distinguishes the two is implausible. Alternatively, such critics urge on conceptual grounds that no clear distinctions can be drawn in these matters, that foreseeing with certainty is indistinguishable from intending (Bennett 1981), that omitting is one kind of causing (Schaffer 2012), and so forth.

Thirdly, there is the worry about "avoision." By casting our categorical obligations in such agent-centered terms, one invites a kind of manipulation that is legalistic and Jesuitical, what Leo Katz dubs "avoision" (Katz 1996). Some think, for example, that one can transform a prohibited intention into a permissible predictive belief (and thus escape intention-focused forms of agent-relative duty) by the simple expedient of finding some other end with which to motivate the action in question. (This form of avoision rather controversially assumes that we can will the reasons for which we act no less than we can will whether we act.)

Such criticisms of the agent-centered view of deontology drive most who accept their force away from deontology entirely and to some form of consequentialism. Alternatively, some of such critics are driven to patient-centered deontology, which we discuss immediately below. Yet still other of such critics attempt to articulate yet a fourth form of agent-centered deontology. This might be called the "control theory of agency." On this view, our agency is invoked whenever our choices could have made a difference. This cuts across the intention/foresight, act/omission, and doing/allowing distinctions, because in all cases we *controlled* what happened through our choices (Frey 1995). Yet as an account of deontology, this seems worrisomely broad. It disallows consequentialist justifications whenever: we foresee the death of an innocent; we omit to save, where our saving would have made a difference and we knew it; where we remove a life-saving device, knowing the patient will die. If deontological norms are so broad in content as to cover all these foreseeings, omittings, and allowings, then good consequences (such as a net saving of innocent lives) are ineligible to justify them. This makes for a wildly counterintuitive deontology: surely I can, for example, justify not throwing the rope to one (and thus omit to save him) in order to save two others equally in need. This breadth of obligation also makes for a conflict-ridden deontology: by refusing to cabin our categorical obligations by the distinctions of the Doctrine of Double Effect and the Doctrine of Doing and Allowing, situations of conflict between our stringent obligations proliferate in a troublesome way (Anscombe 1962).

2.2 Patient-Centered Deontological Theories

A second group of deontological moral theories can be classified as *patient*-centered, not *agent*-centered. These theories are rights-based rather than duty-based; and some versions purport to be quite agent-neutral in the reasons they give moral agents.

All patient-centered deontological theories are properly characterized as theories premised on people's rights. An influential version posits, as its core right (often described by "the Means Principle"), the right against being used only as means for producing good consequences without one's consent. Such a core right is not to be confused with more discrete rights, such as the right against being killed, or being killed intentionally. It is a right against being used by another for the user's or others' benefit. More

specifically, this version of patient-centered deontological theories proscribes the *using* of another's body, labor, and talent without the latter's consent. One finds this notion expressed, albeit in different ways, in the work of the so-called Right Libertarians (e.g., Nozick 1974, Mack 2000), but also in the works of the Left-Libertarians as well (e.g., Steiner 1994; Vallentyne and Steiner 2000; Vallentyne, Steiner, and Otsuka 2005). On this view, the scope of strong moral duties—those that are the correlatives of others' rights—is jurisdictionally limited and does not extend to resources for producing the Good that would not exist in the absence of those intruded upon—that is, their bodies, labors, and talents. In addition to the Libertarians, others whose views include this prohibition on using others include Quinn, Kamm, Alexander, Ferzan, Gauthier, and Walen (Quinn 1989; Kamm 1996; Alexander 2016; Alexander and Ferzan 2009, 2012; Gauthier 1986; Walen 2014, 2016).

Just as do agent-centered theories, so too do patient-centered theories (such as that forbidding the using of another) seek to explain common intuitions about such classic hypothetical cases as Trolley, Transplant, and Fat Man (Thomson 1985). In Trolley, a runaway trolley will kill five workers unless diverted to a siding where it will kill one worker. Most people regard it as permissible and perhaps mandatory to switch the trolley to the siding. By contrast, in Transplant, where a surgeon can kill one healthy patient and transplant his organs to five dying patients, thereby saving their lives, the universal reaction is condemnation. (The same is by-and-large true in Fat Man, where the runaway trolley cannot be switched off the main track but can be stopped before reaching the five workers by pushing a fat man into its path, resulting in his death.)

The injunction against using arguably accounts for these contrasting reactions. After all, in each example, one life is sacrificed to save five. Yet there appears to be a difference in the means through which the net four lives are saved. In Transplant (and Fat Man), the doomed person is used to benefit the others. They could not be saved in the absence of his body. In Trolley, on the other hand, the doomed victim is not used. The workers would be saved whether or not he is present on the second track.

Notice, too, that this patient-centered libertarian version of deontology handles Trolley, Transplant et al. differently from how they are handled by agent-centered versions. The latter focus on the agent's mental state or on whether the agent acted or caused the victim's harm. The patient-centered theory focuses instead on whether the victim's body, labor, or talents were the means by which the justifying results were produced. So one who realizes that by switching the trolley he can save five trapped workers and place only one in mortal danger—and that the danger to the latter is not the means by which the former will be saved—acts permissibly on the patient-centered view if he switches the trolley *even if he does so with the intention of killing the one worker*. Switching the trolley is causally sufficient to bring about the consequences that justify the act—the saving of net four workers—*and it is so even in the absence of the one worker's body, labor, or talents*. (The five would be saved if the one escaped, was never on the track, or did not exist.) By contrast, on the intent and intended action versions of agent-centered theories (but not on the cause-based versions), the one who switches the trolley does *not* act permissibly if he acts with the intention to harm the one worker. (This could be the case, for example, when the one who switches the trolley does so to kill the one whom he hates, only knowing that he will thereby save the other five workmen.) On the patient-centered version, if an act is otherwise morally justifiable by virtue of its balance of good and bad consequences, and the good consequences are achieved without the necessity of using anyone's body, labor, or talents without that person's consent as the means by which they are achieved, then it is morally immaterial (to the *permissibility* of the act but not to the *culpability* of the actor) whether someone undertakes that act with the intention to achieve its bad consequences. (This is true, of course, only so long as the concept of *using*

does not implicitly refer to the intention of the user) (Alexander 2016). And in assessing the culpability of risky conduct, any good consequences must be discounted, not only by the perceived risk that they will not occur, but also by the perceived risk that they will be brought about by a using; for any such consequences, however good they otherwise are, cannot be considered in determining the permissibility and, derivatively, the culpability of acts (Alexander 2016).

Patient-centered deontologists handle differently other stock examples of the agent-centered deontologist. Take the acceleration cases as an example. When all will die in a lifeboat unless one is killed and eaten; when Siamese twins are conjoined such that both will die unless the organs of one are given to the other via an operation that kills the first; when all of a group of soldiers will die unless the body of one is used to hold down the enemy barbed wire, allowing the rest to save themselves; when a group of villagers will all be shot by a blood-thirsty tyrant unless they select one of their numbers to slake the tyrants lust for death—in all such cases, the causing/accelerating-distinguishing agent-centered deontologists would permit the killing but the usings-focused patient-centered deontologist would not.

As another example, consider that of Loop Trolley (Thomson, 1985). Here, as in Trolley, a runaway trolley is headed towards five trapped workmen who will be killed unless bystander at the switch switches the trolley to a second track where there is only one trapped workman. Unlike Trolley, in Loop Trolley the second track loops around to rejoin the first track, meaning that unless the body of the one trapped workman is on the second track to stop the trolley, the five workmen would still be killed even if the trolley is turned onto the second track. According to the patent-centered deontology we are considering, the switchman now may *not* switch the trolley to the second track because to do so would be to "use" the first workman in saving the five (this, because now the body of the one

is needed to save the five); whereas according to the agent-relative version of deontology, by contrast, one may still redirect the runaway trolley.

The restriction of deontological duties to *usings* of another raises a sticky problem for those patient-centered deontological theories that are based on the core right against using: how can they account for the *prima facie* wrongs of killing, injuring, and so forth when done not to use others as means, but for some other purpose or for no purpose at all? The answer is that such patient-centered deontological constraints must be supplemented by consequentialist-derived moral norms to give an adequate account of morality. Killing, injuring, and so forth will usually be unjustifiable on a consequentialist calculus, especially if everyone's interests are given equal regard. It is when killing and injuring are otherwise justifiable that the deontological constraint against using has its normative bite over and against what is already prohibited by consequentialism. (This narrowness of patient-centered deontology makes it counterintuitive to agent-centered deontologists, who regard prohibitions on killing of the innocent, etc., as paradigmatically deontological.)

The patient-centered version of deontology is aptly labeled libertarian in that it is not plausible to conceive of not being aided as being used by the one not aiding. Using is an action, not a failure to act. More generally, it is counterintuitive to many to think that any of us have a *right* to be aided. For if there were a strong (that is, enforceable or coercible) duty to aid others, such that, for example, A had a duty to aid X, Y, and Z; and if A could more effectively aid X, Y, and Z by coercing B and C to aid them (as is *their* duty), then A would have a duty to "use" B and C in this way. For these reasons, any positive duties will not be rights-based ones on the view here considered; they will be consequentially-justified duties that can be trumped by the right not to be coerced to perform them.

Patient-centered deontological theories are often conceived in agent-neutral reason-giving terms. John has a right to the exclusive use of his body, labor, and talents, and such a right gives everyone equal reason to do actions respecting it. But this aspect of patient-centered deontological theories gives rise to the virulent form of the so-called paradox of deontology (Scheffler 1988; Heuer 2011)—that if respecting Mary's and Susan's rights is as important morally as is protecting John's rights, then why isn't violating John's rights permissible (or even obligatory) when doing so is necessary to protect Mary's and Susan's rights from being violated by others? If all rights are equally valuable, why isn't it moral to violate one when doing so prevents the violation of two? Patient-centered deontological theories might arguably do better if they abandoned their pretense of being agent-neutral. They could conceive of rights as giving agent-relative reasons to each actor to refrain from doing actions violative of such rights. Take the core right against being used without one's consent hypothesized earlier. The correlative duty is not to use another without his consent. If such duty is agent-relative, then the rights-based deontologist (no less than the agent-centered deontologist) has the conceptual resources to answer the virulent form of the paradox of deontology. That is, each of us may not use John, even when such using of John would minimize usings of John by others in the future. Such duties are personal to each of us in that we may not justify our violating such a duty now by preventing others' similar violations in the future. Such personal duties are agent-centered in the sense that the agency of each person is central to the duties of each person, so that *your* using of another now cannot be traded off against other possible usings at other times by other people. Patient-centered deontologies are thus arguably better construed to be agent-relative in the reasons they give.

2.3 Contractualist Deontological Theories

Somewhat orthogonal to the distinction between agent-centered versus patient-centered deontological theories are contractualist deontological theories. Morally wrong acts are, on such accounts, those acts that would be forbidden by principles that people in a suitably described social contract would accept (e.g., Rawls 1971; Gauthier 1986), or that would be forbidden only by principles that such people could not "reasonably reject" (Scanlon 2003).

In deontology, as elsewhere in ethics, it is not entirely clear whether a contractualist account is really normative as opposed to metaethical. If such account is a first order normative account, it is probably best construed as a patient-centered deontology; for the central obligation would be to do onto others only that to which they have consented. But so construed, modern contractualist accounts would share the problems that have long bedeviled historical social contract theories: how plausible is it that the "moral magic" of consent is the first principle of morality? And how much of what is commonly regarded as permissible to do to people can (in any realistic sense of the word) be said to be actually consented to by them, expressly or even implicitly?

In fact modern contractualisms look meta-ethical, and not normative. Thomas Scanlon's contractualism, for example, which posits at its core those norms of action that we can justify to each other, is best construed as an ontological and epistemological account of moral notions. The same may be said of David Gauthier's contractualism. Yet so construed, metaethical contractualism as a method for deriving moral norms does not necessarily lead to deontology as a first order ethics. John Harsanyi, for example, argues that parties to the social contract would choose utilitarianism over the principles John Rawls argues would be chosen (Harsanyi 1973). Nor is it clear that meta-ethical contractualism, when it

does generate a deontological ethic, favors either an agent centered or a patient centered version of such an ethic.

2.4 Deontological Theories and Kant

If any philosopher is regarded as central to deontological moral theories, it is surely Immanuel Kant. Indeed, each of the branches of deontological ethics—the agent-centered, the patient-centered, and the contractualist—can lay claim to being Kantian. The agent-centered deontologist can cite Kant's locating the moral quality of acts in the principles or maxims on which the agent acts and not primarily in those acts' effects on others. For Kant, the only thing unqualifiedly good is a good will (Kant 1785). The patient-centered deontologist can, of course, cite Kant's injunction against using others as mere means to one's end (Kant 1785). And the contractualist can cite, as Kant's contractualist element, Kant's insistence that the maxims on which one acts be capable of being willed as a universal law—willed by all rational agents (Kant 1785). (See generally the entry on Kant.)

3. The Advantages and the Disadvantages of Deontological Theories

Having canvassed the two main types of deontological theories (together with a contractualist variation of each), it is time to assess deontological morality more generally. On the one hand, deontological morality, in contrast to consequentialism, leaves space for agents to give special concern to their families, friends, and projects. At least that is so if the deontological morality contains no strong duty of general beneficence, or, if it does, it places a cap on that duty's demands. Deontological morality, therefore, avoids the overly demanding and alienating aspects of consequentialism and accords more with conventional notions of our moral duties.

Likewise, deontological moralities, unlike most views of consequentialism, leave space for the supererogatory. A deontologist can do more that is morally praiseworthy than morality demands. A consequentialist cannot, assuming none of the consequentialists' defensive maneuvers earlier referenced work. For such a pure or simple consequentialist, if one's act is not morally demanded, it is morally wrong and forbidden. Whereas for the deontologist, there are acts that are neither morally wrong nor demanded, some—but only some—of which are morally praiseworthy.

As we have seen, deontological theories all possess the strong advantage of being able to account for strong, widely shared moral intuitions about our duties better than can consequentialism. The contrasting reactions to Trolley, Fat Man, Transplant, and other examples earlier given, are illustrative of this.

Finally, deontological theories, unlike consequentialist ones, have the potential for explaining why certain people have moral standing to complain about and hold to account those who breach moral duties. For the moral duties typically thought to be deontological in character—unlike, say, duties regarding the environment—are duties *to* particular people, not duties to bring about states of affairs that no particular person has an individual right to have realized.

On the other hand, deontological theories have their own weak spots. We here consider three of the most prominent. First, and perhaps the most prominent, is the paradox of deontology mentioned before with respect to patient-centered versions of deontology. The virulent form of the paradox—claiming conceptual incoherence for agent-neutral versions of patient centered deontology—is not present in most forms of deontology. Yet the moral form of the paradox is present for all forms of deontology. It is this: if violations of categorical norms are wrong (or bad), then are not more

such violations—whether by others or even by one's future self (Johnson, 2019)— more wrong (or worse) than fewer? If so, then in situations where one wrongful act would prevent many more such acts, do not the deontological norms prohibiting such acts require us to act in ways that make the world morally worse? And is that not extremely paradoxical?

One way of avoiding this paradox is to deny that the wrongness of acts violative of deontological norms can be translated into the badness of states of affairs (namely, the states of affairs constituted by such acts having been done). Then two wrong acts would not be worse than one, and, more generally, such wrongs could not be summed into anything of normative significance. Such a response would reject the comparability of states of affairs that involve violations and those that do not, as well as the comparability of states of affairs that involve more or fewer violations (Brook 2007). Such a response might attempt to back these assertions by relying on the often-mentioned "separateness of persons": wrongs are always wrongs to some person, and a wrong done to Y and a wrong done to Z cannot be added because there is no person who suffers this amalgamated wrong (cf. Taurek 1977). Whatever the attractions of this solution to the paradox of deontology, they come at a high cost. In Trolley, for example, one could not easily claim that it is not only permissible but better to switch and save five (as is the common intuition). For if wrongs to persons cannot be summed (because there is no person who is the victim of the aggregated wrong), then harms to persons cannot be summed either (for there is no person who suffers the aggregated harm); five deaths is thus not worse than one death, and there is thus no reason to switch to save the five. (Tenenbaum 2023) Worse yet, were the trolley headed for the one worker rather than the five there would be no reason not to switch the trolley to kill the five rather than the one. If the numbers don't count, they seemingly don't count either way. For reasons such as these, few deontologists rely on Taurek's counterintuitive conclusion that the numbers don't count. In situations generally (not just in Trolley) where the

choice is to save a few or save many, many deontologists would save the many so as to acknowledge the importance of each of the extra persons saved by the choice; but other deontologists would flip a coin, use a weighted coin flip, or regard either choice as equally acceptable (Broome 1998; Doggett 2013; Doucet 2013; Dougherty 2013; Halstead 2016; Henning 2015; Hirose 2007; Hsieh et al. 2006; Huseby 2011; Kamm 1993; Rasmussen 2012; Sauders 2009; Scanlon 2003; Suikkanen 2004; Timmerman 2004; Wasserman and Strudler 2003). The calculation becomes even more complicated in situations where there is a difference between the types of harm caused from the type of harm prevented, as when causing the death of one innocent is the only way to prevent a million people to have a headache (Tomlin 2017; Tadros 2019; Horton 2018, Steuwer 2022). The calculation becomes further complicated yet if the entities to whom deontological duties are owed or who hold rights are themselves of different moral status (Kagan 2019).

Going the other way to deal with the moral version of the paradox of deontology has its own problems. That is, suppose one allows there to be some translation of the wrongness of violating a deontological norm into the badness of the state of affairs constituted by such a wrong having been done; then the reason not to do some action violative of a deontological prohibition—even when doing so will minimize the future doings of that action—will have to compete with the reason based on the maximization of good states of affairs. This raises two questions for the deontologist: one, is there some form of rationality, alternative to the maximization of the Good of the consequentialists, that can make sense of pure norm-conformity as generative of reasons for action? And two, how can we be certain that the deontological norm-created reason for not doing the prohibited action will always win out in competition with that reason in favor of doing the action which is based on the minimization of future violations; after all, deontological obligations come in different stringencies, so why cannot violation of a modestly stringent deontological

prohibition be outweighed by the heavily weighted goodness of preventing numerous future violations? To answer both questions, the deontologist seemingly needs an account like that of Joseph Raz's idea of a "protected reason": the norm itself both has positive reason-giving potential and it negatively strips (Raz: "excludes") what otherwise would be good reasons of their normal reason-giving potential (Raz 1999). Both of these ideas require considerable elaboration and defense to be plausible (Moore 2025).

Both of these routes to solving the paradox of deontology are thus problematic. The problem of how to account for the significance of numbers without giving up deontology by adopting consequentialism, and without resurrecting the paradox of deontology, is one that a number of deontologists are now working to solve(e.g., Kamm 1996; Scanlon 2003; Otsuka 2006; Hsieh et al. 2006; Steuwer 2023). Until it is solved, it will remain a huge thorn in the deontologist's side.

A second crucial problem with which deontologists must grapple is the problem posed by the conflicts that seem to exist between various deontological obligations. (Please see entry on moral dilemmas.) It would be a widely unfair morality—a kind of cruel joke on the human race—as well as a violation of at least *standard* deontic logic, for morality to admit of the possibility that an actor would be simultaneously obligated both to do and to refrain from doing some particular act. Kant's bold proclamation that "a conflict of duties … is inconceivable" (Kant 1797 [1965, p. 25; AK 6: 224]) is the conclusion wanted, but reasons for believing it are difficult to produce for those who hold that morality consists of categorical norms.

One of the advantages claimed for agent-centered versions of deontology is that the various agency-determining variables restrict the scope of deontological obligations in ways that lessen the possibilities of there being conflicts of deontological duties (Anscombe 1962). In Trolley, for example, the switchman is not seen to face a conflict of two deontological duties (a duty not to kill one workman versus a duty not to kill five workmen); rather, if he does not turn the trolley he is only omitting to save the five, not causing their death, and if he turns the trolley he is only (redirectively) allowing the death of the one, not causing that death, so on both sides of the balance he acts outside of the scope of deontological obligations (this is what makes him free to maximize good consequences). Still, while such conflict-reducing possibilities are real, few think it likely that they can be sufficient to eliminate conflict entirely. Moreover, the more restrictive of scope such agency limitations are made to be, the more they invite the kind of manipulation and evasion of deontology's norms that Leo Katz calls, "avoision" (Katz 1996).

That leaves two alternative ways for the deontologist to reach the desired Kantian conclusion that there can b e no conflict of deontological duties. One is to go the route of the specificationist: individuate the norms of deontology so finely that the content of each contains priority rules and exceptions that resolve any apparent conflict with all other such norms (see the entry on Rights). This makes such apparent staples of deontology as the norm, "Do not murder innocents," into a mere epistemic aid that is abbreviatory of much longer contents of true moral norms; what is abbreviated in the true moral norm is something like, "don't murder innocents unless...[they are inadvertently going to kill you, they are in extreme pain and request painless death, etc., etc.]." Many find such specificationism implausible because it is (seemingly) so after-the-fact and *ad hoc*. Others, however, find such specificationism to be inevitable, not only because of its elimination of conflict between duties but also because: (1) it is demanded by the fine grained individuation of the kinds of moral wrongs a properly proportionate retributive punishment scheme requires, finely individuated wrongs requiring equally finely individuated norms whose violation constitutes such wrongs (Hurd and Moore 2021); and (2) because it lessens the need of deontology—a need whose satisfaction is

otherwise puzzling to non-theists (Williams 1973)—for there to be some general text of moral laws seemingly requiring an authorative law-giver as its source.

The other well known approach to deal with the possibility of conflict between deontological duties is to reduce the force of such duties to that of only "*prima facie*" duties (Ross 1930, 1939). Then one can claim that all apparent conflicts of duties are in reality conflicts between merely *prima facie* duties, which is unproblematic so long as it does not infect what one is categorically obligated to do as one's concrete, over-all duty on some occasion (Tohaneanu 2015). The adequacy of this line of response depends mightily on two variables. One is what reasons count in overriding a prima facie duty so that there is no corresponding all-out duty: are these considerations limited to the reasons constituted by other deontological norms, or do they include the reasons constituted by good states of affairs? The second variable is usually put as the difference between an epistemic reading of "*prima facie*" and the substantive (or "pro tanto") reading. On the first reading there is no independent force to a *prima facie* duty, it being only an epistemic aid to finding what does have force which is the all-out duty; on the second reading the *prima facie* duty does have force independently of its corresponding all-out duty, but it is not the force of obligating action but only of apologizing, compensating, etc., for not doing the action one was prima facie (but not all-out) obligated to do.

These two variables chart the dangers and limitations of the "*prima facie* duties" retort to the problem of conflicting duties. If the existence of good states of affairs can constitute the kind of reason that can block a prima facie duty from being an all-out duty, then deontology so construed is in danger of collapsing into consequentialism. This is particularly true if "prima facie" is read epistemically—for then the norms of deontology are mere epistemic aids to counting up what really counts, which are good states of affairs. But it is also true of the pro tanto reading (because even there the deontological bottom line about what act one must do loses out to good consequences). If the reasons that can block a prima facie duty from being an all-out duty do not include good states of affairs but are limited to other deontological norms, then the prima facie duties response to conflict is in danger of collapsing into the specificationist response. This is particularly obvious on the epistemic reading of "*prima facie*"; for then the all-out duty's content wholly depends on how the content of all "*prima facie*" duties interlock with one another via exceptions and priority rules. But this is also true for the pro tanto reading of "*prima facie*." For again, what one is all-out obligated to do (as opposed to apologize for not doing, etc.) includes the exceptions and priorities that make the allowances needed to avoid conflict.

A third set of problems for deontology arise from the fact that there are situations—unfortunately not all of them thought experiments—where compliance with deontological norms will bring about disastrous consequences (Zack 2023). To take a stock example of much recent public discussion, suppose that unless *A* violates the deontological duty not to torture an innocent person (*B*), ten, or a thousand, or a million other innocent people will die because of a hidden nuclear device. If *A* is forbidden by deontological morality from torturing *B*, many would regard that as a *reductio ad absurdum* of deontology.

Deontologists have six possible ways of dealing with what Robert Nozick called, "moral catastrophes" (Nozick 1974), although only two of these are very plausible. First, they can just bite the bullet and declare that sometimes doing what is morally right will have tragic results but that allowing such tragic results to occur is still the right thing to do. Complying with moral norms will surely be difficult on those occasions, but the moral norms apply nonetheless with full force, overriding all other considerations. We might call this the Kantian response, after Kant's

endorsement of the saying "*fiat iustitia, pereat mundus*" ["do justice though the world perish"] (Kant 1795 [1996, 345]; AK 8:378). One might also call this the absolutist conception of deontology, because such a view maintains that conformity to norms has absolute force and not merely great weight (Waldron 2012; Moore and Waldron 2022).

The second plausible response is for the deontologist to abandon Kantian absolutism for what is usually called "threshold" or "moderate" deontology. A threshold deontologist holds that deontological norms govern up to a point despite adverse consequences; but when the consequences become so dire that they cross the stipulated threshold, consequentialism takes over (Moore 1997, ch. 17). *A* may not torture *B* to save the lives of two others, but he may do so to save a thousand lives if the "threshold" is higher than two lives but lower than a thousand.

There are quite a few varieties of threshold deontology (Kagan 2019) but two that are particularly worth distinguishing. On the simple version, there is some fixed threshold of awfulness beyond which morality's categorical norms no longer have their overriding force. Such a threshold is fixed in the sense that it does not vary with the stringency of the categorical duty being violated. The alternative is what might be called "sliding scale threshold deontology." On this version, the threshold varies in proportion to the degree of wrong being done—the wrongness of stepping on a snail has a lower threshold (over which the wrong can be justified) than does the wrong of stepping on a baby.

There are also two ethical contexts for application of threshold deontology that are worth distinguishing. Historically the context of discussion garnering the most attention has been the override of obligation done by reaching a threshold of awful consequences prevented by doing what the obligation forbids doing. But equally plausible is to think that the assertion of agent-centered prerogatives also must yield once the assertion of such

prerogatives reaches a threshold of awful consequences by doing what the prerogative permits (Moore 2012; Brennan 2018).

Threshold deontology—in all of its guises and in all of its contexts of application—is an attempt to save deontological morality from the charge of fanaticism in how it deals with moral catastrophes (Similansky 2003; Cook 2018, Johnson, 2020). There is some similarity to the "*prima facie* duty" version of deontology developed to deal with the problem of conflicting duties, yet threshold deontology is usually interpreted with such a high threshold that it more closely mimics the outcomes reached by a "pure," absolutist kind of deontology.

Threshold deontology faces several theoretical difficulties. Foremost among them is giving a theoretically tenable account of the location of such a threshold, either absolutely or on a sliding scale (Alexander 2000; Ellis 1992; Moore 2019; Arneson 2019; Cole 2019; Alexander 2019). Why is the threshold for torture of the innocent at one thousand lives, say, as opposed to nine hundred or two thousand? An aspect of this problem is that whatever the threshold, as the dire consequences approach it, counter-intuitive results appear to follow. For example, it may be permissible, if we are one-life-at-risk short of the threshold, to pull one more person into danger who will then be saved, along with the others at risk, by killing an innocent person (Alexander 2000). Also, there is some uncertainty about how one is to reason after the threshold has been reached: are we to calculate at the margin on straight consequentialist grounds, use an agent-weighted mode of summing, or do something else? A final problem is that threshold deontology threatens to collapse into a kind of consequentialism. Indeed, it can be perhaps shown that the sliding scale version of threshold deontology is extensionally equivalent to an agency-weighted form of consequentialism (Sen 1982).

The remaining four strategies for dealing with the problem of dire consequence cases all have the flavor of evasion by the deontologist. Consider first the famous view of Elizabeth Anscombe: such cases (real or imagined) can never present themselves to the consciousness of a truly moral agent because such agent will realize it is immoral to even think about violating moral norms in order to avert disaster (Anscombe 1958; Geach 1969; Nagel 1979). Such rhetorical excesses should be seen for what they are, a peculiar way of stating Kantian absolutism motivated by an impatience with the question.

Another response by deontologists, this one most famously associated with Bernard Williams, shares some of the "don't think about it" features of the Anscombean response. According to Williams (1973), situations of moral horror are simply "beyond morality," and even beyond reason. (This view is reminiscent of the ancient view of natural necessity, revived by Sir Francis Bacon, that such cases are beyond human law and can only be judged by the natural law of instinct.) Williams tells us that in such cases we just act. Interestingly, Williams contemplates that such "existentialist" decision-making will result in our doing what we have to do in such cases —for example, we torture the innocent to prevent nuclear holocaust.

Surely this is an unhappy view of the power and reach of human law, morality, or reason. Indeed, Williams (like Bacon and Cicero before him) thinks there is an answer to what should be done, albeit an answer very different than Anscombe's. But both views share the weakness of thinking that morality and even reason runs out on us when the going gets tough.

Yet another strategy is to divorce completely the moral appraisals of acts from the blameworthiness or praiseworthiness of the agents who undertake them, even when those agents are fully cognizant of the moral appraisals. So, for example, if *A* tortures innocent *B* to save a thousand others, one can hold that *A*'s act is morally wrong but also that *A* is morally praiseworthy for having done it.

Deontology does have to grapple with how to mesh deontic judgments of wrongness with "hypological" (Zimmerman 2002) judgments of blameworthiness (Alexander 2004). Yet it would be an oddly cohering morality that condemned an act as wrong yet praised the doer of it. Deontic and hypological judgments ought to have more to do with each other than that. Moreover, it is unclear what action-guiding potential such an oddly cohered morality would have: should an agent facing such a choice avoid doing wrong, or should he go for the praise?

The last possible strategy for the deontologist in order to deal with dire consequences is to distinguish moral reasons from all-things-considered reasons and to argue that whereas moral reasons dictate obedience to deontological norms even at the cost of catastrophic consequences, all-things-considered reasons dictate otherwise. (This is one reading of Bernard William's famous discussion of moral luck, where non-moral reasons seemingly can trump moral reasons (Williams 1975, 1981); this is also a strategy some consequentialists (e.g., Portmore 2003) seize as well in order to handle the demandingness and alienation problems endemic to consequentialism.) But like the preceding strategy, this one seems desperate. Why should one even care that moral reasons align with deontology if the important reasons (namely, the all-things-considered reasons that actually govern decisions) align with consequentialism?

4. Deontology's Relation(s) to Consequentialism Reconsidered

The perceived weaknesses of deontological theories have led some to consider how to eliminate or at least reduce those weaknesses while preserving deontology's advantages. One way to do this is to embrace both

consequentialism and deontology, combining them into some kind of a mixed theory. Given the differing notions of rationality underlying each kind of theory, this is easier said than done. After all, one cannot simply weigh agent-relative reasons against agent-neutral reasons, without stripping the former sorts of reasons of their distinctive character.

A time-honored way of reconciling opposing theories is to allocate them to different jurisdictions. Tom Nagel's reconciliation of the two theories is a version of this, inasmuch as he allocates the agent-neutral reasons of consequentialism to our "objective" viewpoint, whereas the agent-relative reasons of deontology are seen as part of our inherent subjectivity (Nagel 1986). Yet Nagel's allocations are non-exclusive; the same situation can be seen from either subjective or objective viewpoints, meaning that it is mysterious how we are to combine them into some overall view.

A less mysterious way of combining deontology with consequentialism is to assign to each a jurisdiction that is exclusive of the other. One possibility here is to regard the agent-neutral reasons of consequentialism as a kind of default rationality/morality in the sense that when an agent-relative permission or obligation applies, it governs, but in the considerable logical space where neither applies, consequentialism holds sway; agent-relative permissions and obligations in this sense "side-constrain" the maximization of good states of affairs (Nozick, 1974). Remembering that for deontologists who are threshold deontologists, consequentialist reasons may still determine right action even in areas governed by agent-relative obligations or permissions, once the level of bad consequences crosses the relevant threshold.

4.1 Making no concessions to consequentialism: a purely deontological rationality?

In contrast to mixed theories, deontologists who seek to keep their deontology pure hope to expand agent-relative reasons to cover all of morality and yet to mimic the advantages of consequentialism. Doing this holds out the promise of denying sense to the otherwise damning question, how could it be moral to make (or allow) the world to be worse (for they deny that there is any states-of-affairs "worseness" in terms of which to frame such a question) (Foot 1985). To make this plausible, one needs to expand the coverage of agent-relative reasons to cover what is now plausibly a matter of consequentialist reasons, such as positive duties to strangers. Moreover, deontologists taking this route need a content to the permissive and obligating norms of deontology that allows them to mimic the outcomes making consequentialism attractive. This requires a specificationist picture of morality's norms that is extremely detailed in content, so that what looks like a consequentialist balance can be generated by a complex series of norms with extremely detailed priority rules and exception clauses (Richardson 1990). Few consequentialists will believe that this is a viable enterprise.

4.2 Making no concessions to deontology: a purely consequentialist rationality?

The mirror image of the pure deontologist just described is the indirect or two-level consequentialist. For this view too seeks to appropriate the strengths of both deontology and consequentialism, not by embracing both, but by showing that an appropriately defined version of one can do for both. The indirect consequentialist, of course, seeks to do this from the side of consequentialism alone.

Yet as many have argued (Lyons 1965; Alexander 1985), indirect consequentialism collapses either into: blind and irrational rule-worship ("why follow the rules when not doing so produces better consequences?"); direct consequentialism ("acts in conformity to the rules rather miraculously produce better consequences in the long run"); or nonpublicizability ("ordinary folks should be instructed to follow the rules but should not be told of the ultimate consequentialist basis for doing so, lest they depart from the rules mistakenly believing better consequences will result"). For more information, please see the entry on rule consequentialism. Nor can the indirect consequentialist adequately explain why those who violate the indirect consequentialist's rules have "wronged" those who might be harmed as a result, that is, why the latter have a personal complaint against the former. (This is true irrespective of whether the rule-violation produces good consequences; but it is especially so when good consequences result from the rule-violation.) The bottom line is that if deontology has intuitive advantages over consequentialism, it is far from obvious whether those advantages can be captured by moving to indirect consequentialism, even if there is a version of indirect consequentialism that could avoid the dire consequences problem that bedevils deontological theories.

5. Deontology and Uncertainty About Outcomes

Recently, deontologists have begun to ask how an actor should evaluate courses of action in which it is uncertain whether a deontological constraint will be violated. For example, should one detonate dynamite in a mining operation if there is a chance that the explosion will cause the Fat Man to tumble into the path of the trolley that would otherwise kill five? (Assume that were the chance the same that the explosion would instead divert the trolley in Trolley, killing one but saving five, the detonation would be permissible.) Or should one take a drive to observe the scenery if there is a slightly increased chance that, because of the possibility of

traffic, doing so will cause one to miss a lunch one had promised to attend? Whether deontological constraints focus on agents' intentions or beliefs, or whether they focus on agents' counting positively in their deliberations others' use as means, how should the uncertainty of outcomes be taken into account by deontologists? This question has been addressed by Aboodi, Borer, and Enoch (2008); Alexander (2016; 2018); Lazar (2015; 2017a, 2017b, 2018); Smith (2014); Tarsney (2018); Tenenbaum (2017) and Tomlin (2019).

6. Deontological Theories and Metaethics

Deontological theories are normative theories. They do not presuppose any particular position on moral ontology or on moral epistemology. Presumably, a deontologist can be a moral realist of either the natural (moral properties are identical to or grounded in natural properties) or nonnatural (moral properties are not themselves natural properties even if they are nonreductively related to natural properties) variety. Or a deontologist can be an expressivist, a constructivist, a transcendentalist, a conventionalist, or a Divine command theorist regarding the nature of morality. Likewise, a deontologist can claim that we know the content of deontological morality by direct intuition, by Kantian reflection on our normative situation, or by reaching reflective equilibrium between our particular moral judgments and the theories we construct to explain them (theories of intuitions).

Nonetheless, although deontological theories can be agnostic regarding metaethics, some metaethical accounts seem less hospitable than others to deontology. For example, the stock furniture of deontological normative ethics—rights, duties, permissions—fits uneasily in the realist-naturalist's corner of the metaethical universe. (Which is why many naturalists, if they are moral realists in their meta-ethics, are consequentialists in their ethics.) Nonnatural realism, conventionalism, transcendentalism, and Divine

command seem more hospitable metaethical homes for deontology. (For example, the paradox of deontology above discussed may seem more tractable if morality is a matter of personal directives of a Supreme Commander to each of his human subordinates.) If these rough connections hold, then weaknesses with those metaethical accounts most hospitable to deontology will weaken deontology as a normative theory of action. Some deontologists have thus argued that these connections need not hold and that a naturalist-realist meta-ethics can ground a deontological ethics (Moore 2004).

Bibliography

Aboodi, R., A. Borer, and D. Enoch, 2008, "Deontology, Individualism, and Uncertainty: A Reply to Jackson and Smith," *Journal of Philosophy*, 105: 259–272.

Alexander, L., 1985, "Pursuing the Good—Indirectly," *Ethics*, 95(2): 315–332.

——, 2000, "Deontology at the Threshold," *San Diego Law Review*, 37(4): 893–912.

——, 2004, "The Jurisdiction of Justice: Two Conceptions of Political Morality," *San Diego Law Review*, 41(3): 949–966.

——, 2016, "The Means Principle," in K.K. Ferzan and S.J. Morse (eds.), *Legal, Moral, and Metaphysical Truths: The Philosophy of Michael S. Moore*, Oxford: Oxford University Press, 251–264.

——, 2018, "The Need to Attend to Probabilities—For Purposes of Self-Defense and Other Preemptive Actions," *San Diego L. Rev*, 55: 223–229.

——, 2019, "Responses and Appreciations," *Moral Puzzles and Legal Perspectives*, H. Hurd (ed.), Cambridge: Cambridge University Press, 407–440.

Alexander, L. and K.K. Ferzan, 2009, *Crime and Culpability: A Theory of Criminal Law*, Cambridge: Cambridge University Press.

——, 2012, " 'Moore or Less' Causation and Responsibility: Reviewing Michael S. Moore, *Causation and Responsibility: An Essay in Law, Morals and Metaphysics*," *Criminal Law and Philosophy*, 6(1): 81–92.

Anscombe, G.E.M., 1958, "Modern Moral Philosophy," *Philosophy*, 33(124): 1–19, at 10.

——, 1962, "War and Murder," in *Nuclear Weapons: A Catholic Response*, W. Stein (ed.), New York: Sheed and Ward.

Arneson, R., 2019, "Deontology's Travails," *Moral Puzzles and Legal Perspectives*, H. Hurd (ed.), Cambridge: Cambridge University Press, 350–370.

Aquinas, T., 1952, *Summa Theologica*, Chicago: Encyclopedia Britannica.

Bennett, J., 1981, "Morality and Consequences," in *The Tanner Lectures on Human Values* (Volume 2), S. McMurrin (ed.), Cambridge: Cambridge University Press.

Bentham, J., 1789 (1948), *An Introduction to the Principles of Morals and of Legislation*, Oxford: Basil Blackwell.

Bradley, B., 2006, "Against Satisficing Consequentialism," *Utilitas*, 18(2): 97–108.

Brennan, S., 2018, "The Structure of Thresholds for Options," in Mark Timmons (ed.), *Oxford Studies in Normative Ethics*, 8: 192–210.

Brody, B., 1996, "Withdrawing of Treatment Versus Killing of Patients," in *Intending Death*, T. Beauchamp (ed.), Upper Saddle River: Prentice-Hall.

Brook, R., 2007, "Deontology, Paradox, and Moral Evil," *Social Theory and Practice*, 33(3): 431–40.

Broome, J., 1998, "Review: Kamm on Fairness," *Philosophy and Phenomenological Research*, 58 (4): 955–61.

Cole, K., 2019, "Two Cheers for Threshold Deontology," *Moral Puzzles and Legal Perspectives*, H. Hurd (ed.), Cambridge: Cambridge University Press: 388–406.

Cook, T., 2018, "Deontologists Can Be Moderate," *Journal of Value Inquiry*, 52: 199–212.

Doggett, T., 2013, "Saving the Few," *Noûs*, 33: 302–315.

Doucet, M., 2013, "Playing Dice with Morality: Weighted Lotteries and the Number Problem," *Utilitas*, 25: 161–181.

Dougherty, T., 2013, "Rational Numbers: A Non-Consequentialist Explanation of Why You Should Save the Many and Not the Few," *Philosophical Quarterly*, 63: 413–427.

Davis, N., 1984, "The Doctrine of Double Effect: Problems of Interpretation," *Pacific Philosophical Quarterly*, 65: 107–123.

Ellis, A., 1992, "Deontology, Incommensurability and the Arbitrary," *Philosophical and Phenomenological Research*, 52(4): 855–875.

Foot, P., 1967, "The Problem of Abortion and the Doctrine of Double Effect," *Oxford Review*, 5: 5–15.

–––, 1985, "Utilitarianism and the Virtues," *Mind*, 94: 107–123.

Frey, R.G., 1995, "Intention, Foresight, and Killing," in *Intending Death*, T. Beauchamp (ed.), Upper Saddle River: Prentice-Hall.

Gauthier, D., 1986, *Morals By Agreement*, Oxford: Clarendon Press.

Geach, P., 1969, *God and the Soul*, New York: Shocken Books.

Halstead, J., 2016, "The Numbers Always Count," *Ethics*, 126: 789–802.

Harsanyi, J., 1973, *Can the Maximin Principle Serve as a Basis for Morality?: A Critique of John Rawls's Theory*, Berkeley: Center for Research in Management Science.

Hart, H.L.A. and T. Honore, 1985, *Causation in the Law*. New York: Oxford University Press, 2nd edition.

Heuer, U., 2011, "The Paradox of Deontology Revisited," *Oxford Studies in Normative Ethics*, M. Timmons (ed.), Oxford: Oxford University Press, 236–267.

Henning, T., 2015, "From Choice to Chance? Saving People, Fairness, and Lotteries," *Philosophical Review*, 124: 169–206.

Hirose, I., 2007, "Weighted Lotteries in Life and Death Cases," *Ratio*, 20(1): 45–56.

–––, 2015, *Moral Aggregation*, Oxford: Oxford University Press.

Horton, J., 2018, "Always Aggregate," *Philosophy and Public Affairs*, 46(2): 160–174.

Hsieh, N., A. Strudler, and D. Wasserman, 2006, "The Numbers Problem," *Philosophy and Public Affairs*, 34(4): 352–372.

Hurd, H.M., 1994, "What in the World is Wrong?" *Journal of Contemporary Legal Issues*, 5: 157–216.

–––, 1995, "The Deontology of Negligence," *Boston University Law Review*, 76: 249–272.

Hurd, H. and M. Moore, 2021, "The Ethical Implications of Proportioning Punishment to Deontological Desert," *Criminal Law and Philosophy*, 15: 495–514.

Hurka, T., 2019, "More Seriously Wrong, More Importantly Right," *Journal of the American Philosophical Association*, 5(1): 41–58.

Huseby, R., 2011, "Spinning the Wheel or Tossing a Coin?" *Utilitas*, 23(2): 127–39.

Johnson, C., 2019, "The Intrapersonal Paradox of Deontology," *Journal of Moral Philosophy*, 16: 279–301.

–––, 2020, "How Deontologists Can Be Moderate (and Why They Should Be)," *Journal of Value Inquiry*, 154: 1–17.

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

–––, 2019, *How to Count Animals*, Oxford: Oxford University Press.

Kamm, F. M., 1993, *Morality, Morality: Volume I: Death and Whom to Save From It*, New York: Oxford University Press.

–––, 1994, "Action, Omission, and the Stringency of Duties," *University of Pennsylvania Law Review*, 142(5): 1493–1512.

–––, 1996, *Morality, Mortality: Volume II: Rights, Duties and Status*, New York: Oxford University Press.

–––, 2007, *Intricate Ethics: Rights, Responsibilities, and Permissible Harms*, Oxford: Oxford University Press.

Kant, I., 1785, *Groundwork of the Metaphysic of Morals*, H.J. Paton (trans.), New York: Harper and Row, 1964.

–––, 1795 [1996], *Toward Perpetual Peace: A Philosophical Project*, Konigsberg: F. Nicolovius. AK 8: 341–86, translated and edited by Mary Gregor in *Practical Philosophy*, Cambridge: Cambridge University Press, pp. 311–51.

–––, 1797 [1965], *The Metaphysical Elements of Justice: Part I of the Metaphysics of Morals*, J. Ladd (trans.), Indianapolis: Hackett Publishing Co.

Katz, L., 1996, *Ill-Gotten Gains: Evasion, Blackmail, Fraud and Kindred Puzzles of the Law*, Chicago: University of Chicago Press.

Lazar, S., 2015, "Risky Killing and the Ethics of War," *Ethics*, 126: 91–117.

–––, 2017a, "Risky Killing: How Risks Worsen Violations of Objective Rights," *Journal of Moral Philosophy*, 16: 1–26.

–––, 2017b, "Deontological Decision Theory and Agent-Centered Options," *Ethics*, 127: 579–609.

–––, 2018, "In Dubious Battle: Uncertainty and the Ethics of Killing," *Philosophical Studies*, 175: 859–883.

Lyons, D., 1965, *The Forms and Limits of Utilitarianism*, Oxford: Oxford University Press.

Mack, E., 2000, "In Defense of the Jurisdiction Theory of Rights," *Journal of Ethics*, 4: 71–98.

MacMahan, J., 2003, *The Ethics of Killing*, Oxford: Oxford University Press.

Moore, M., 1993, *Act and Crime: The Implications of the Philosophy of Action for the Criminal Law*, Oxford: Clarendon Press.

–––, 1995, "Good Without God," *Natural Law, Liberalism, and Morality*, R. George, ed., Oxford: Oxford University Press, 221–270.

–––, 1997, *Placing Blame: A General Theory of the Criminal Law*, Oxford: Oxford University Press.

–––, 2004, *Objectivity in Ethics and Law*, Aldershot: Ashgate.

–––, 2008, "Patrolling the Borders of Consequentialist Justifications: The Scope of Agent-Relative Obligations," *Law and Philosophy*, 27(1): 35–96.

–––, 2012, "Ethics in Extremis: Targeted Killings and the Morality of Targeted Killings," in *Targeted Killing: Law and Morality in an Asymmetrical World*, C. Finkelstein, J. Ohlin, and Al Altman, (eds.), Oxford: Oxford University Press.

–––, 2019, "The Rationality of Threshold Deontology," *Moral Puzzles and Legal Perspectives*, H. Hurd (ed.), Cambridge: Cambridge University Press, 371–387.

–––, 2025, "Requiem for a Concept: Exclusionary Reasons," in *Engaging Joseph Raz: Themes in Normative Philosophy*, K. Brownlee, D. Enoch, and A. Marmor, (ed.), Oxford: Oxford University Press.

Moore, M., and Hurd, H.M. 2011, "Blaming the Stupid, Clumsy, Selfish, and Weak: The Culpability of Negligence," *Criminal Law and Philosophy*, 5(2): 147–198.

Moore, M., and Waldron, J., 2022, "How Absolute Are Moral Absolutes?," in *Death and Anti-Death Vol. 19: One Year After Judith Jarvis Thomson*, Charles Tandy, (ed.), Ann Arbor: Ria University Press, 189–222.

Nagel, T., 1979, "War and Massacre," in *Mortal Questions*, Cambridge: Cambridge University Press.

–––, 1986, *The View from Nowhere*, New York: Oxford University Press.

Nozick, R., 1974, *Anarchy, State and Utopia*, New York: Basic Books.

Otsuka, M., 2006, "Saving Lives, Moral Theories and the Claims of Individuals," *Philosophy and Public Affairs*, 34(2): 109–135.

Parfit, D., 1984, *Reasons and Persons*, Oxford: Clarendon Press.

Portmore, D.W., 2003, "Position-Relative Consequentialism, Agent-Centered Options, and Supererogation," *Ethics*, 113(2): 303–332.

Quinn, W.S., 1989, "Actions, Intentions, and Consequences: The Doctrine of Doing and Allowing," *The Philosophical Review*, 98(3): 287–312.

Quinton, A., 2007, *Utilitarian Ethics*, 2nd edition, London: Duckworth, pp. 2–3.

Rachels, J., 1975, "Active and Passive Euthanasia," *New England Journal Of Medicine*, 292(2): 78–80.

Rasmussen, K.B., 2012, "Should the Probabilities Count?," *Philosophical Studies*, 159: 205–218.

Rawls, J., 1971, *A Theory of Justice*, Cambridge: Belknap Press of Harvard University Press.

Raz, J., 1999, *Practical Reasons and Norms*, 2nd edition, Oxford: Oxford University Press.

Richardson, H.S., 1990, "Specifying Norms as a Way to Resolve Concrete Ethical Problems," *Philosophy and Public Affairs*, 19(4): 279–310.

Ross, W.D., 1930, *The Right and the Good*, Oxford: Oxford University Press.

–––, 1939, *The Foundations of Ethics*, Oxford: Oxford University Press.

Saunders, B., 2009, "A Defence of Weighted Lotteries in Life Saving Cases," *Ethical Theory and Moral Practice*, 12(3): 279–90.

Scanlon, T.M., 2003, *The Difficulty of Tolerance: Essays in Political Philosophy*, Cambridge: Cambridge University Press.

Schaffer, J., 2012, "Disconnection and Responsibility," *Legal Theory*, 18: 399–435.

Scheffler, S., 1982, *The Rejection of Consequentialism*, Oxford: Oxford University Press.

––– (ed.), 1988, "Introduction," in *Consequentialism and Its Critics*, Oxford: Oxford University Press.

Sen, A.K., 1982, "Rights and Agency," *Philosophy and Public Affairs*, 11(1): 3–39.

Slote, M.A., 1984, *Common-Sense Morality and Consequentialism*, London: Routledge & Kegan Paul.

Smilansky, S., 2003, "Can Deontologists Be Moderate?," *Utilitas*, 15: 71–75.

Smith, H.M., 2014, "The Subjective Moral Duty to Inform Oneself Before Acting to Inform Oneself Before Acting," *Ethics*, 125(1): 11–38.

Steiner, H., 1994, *An Essay on Rights*, Oxford: Blackwell.

Steuwer, B., 2022, "Aggregation, Balancing, and Respect for the Claims of Individuals," *Utilitas*, 33(1): 17–34.

–––, 2023, "Constraints, You and Your Victims," *Noûs*, 57: 942–957.

Suikkanen, J., 2004, "What We Owe to Many," *Social Theory and Practice*, 30(4): 485–506.

Tadros, V., 2019, "Localized Restricted Aggregation," D. Sobel (ed.) *Oxford Studies in Political Philosophy*, Oxford: Oxford University Press.

Tarsney, C., 2018, "Moral Uncertainty for Deontologists," *Ethical Theory and Moral Practice*, 21: 505–520.

Taurek, J.M., 1977, "Should the Numbers Count?" *Philosophy and Public Affairs*, 6(4): 293–316.

Tenenbaum, S., 2017, "Action, Deontology, and Risk: Against the Multiplicative Model," *Ethics*, 127(3): 674–707.

–––, 2023, "Can't Kant Count? Innumerate Views on Saving the Many Over Saving the Few," M. Timmons (ed.), *Oxford Studies in Normative Ethics*, Oxford: Oxford University Press.

Thomson, J.J., 1985, "The Trolley Problem," *Yale Law Journal*, 94: 1395–1415.

Timmerman, J., 2004, "The Individualist Lottery: How People Count, but Not Their Numbers," *Analysis*, 64(2): 106–112.

Tohaneanu, C., 2015, "Could Ross' Pluralist Deontology Solve the Conflicting Duties Problem?," *Revue Roumaine de Philosophie*, 59: 109–124.

Tomlin, P., 2017, "On Limited Aggregation," *Philosophy and Public Affairs*, 45 (3): 232–260.

——, 2019, "Subjective Proportionality," *Ethics*, 129: 254–283.

Vallentyne, P. and H. Steiner (eds.), 2000, *Left-Libertarianism and Its Critics*, Houndmills: Palgrave.

Vallentyne, P., H. Steiner, and M. Otsuka, 2005, "Why Left-Libertarianism Is Not Incoherent, Indeterminate, or Irrelevant: A Reply to Fried," *Philosophy and Public Affairs*, 33(2): 201–215.

Waldron, J., 2012, "What Are Moral Absolutes Like?," *The Harvard Review of Philosophy*, 18: 4–30.

Walen, A., 2014, "Transcending the Means Principle," *Law & Philosophy*, 33: 427–464.

——, 2016, "The Restricting Claims Principle Revisited: Grounding the Means Principle on the Agent-Patient Divide," *Law & Philosophy*, 35: 211–247.

Wasserman, D. and A. Strudler, 2003, "Can a Nonconsequentialist Count Lives?" *Philosophy and Public Affairs*, 31(1): 71–94.

Wedgewood, R., 2011, "Defending Double Effect," *Ratio*, 24(4): 384–401.

Williams, B., 1973, "A Critique of Utilitarianism" in *Utilitarianism: For and Against*, J.J.C. Smart and B. Williams, Cambridge: Cambridge University Press, 77–150.

——, 1975, "Moral Luck," *Proceedings of the Aristotelian Society*, 50(supp.): 115–135; revised and reprinted in Williams 1981.

——, 1981, *Moral Luck*, Cambridge: Cambridge University Press.

Williams, G.L., 1961, *Criminal Law: The General Part*, London: Stevens and Sons, 2nd edition, p. 739.

Woodward, P.A. (ed.), 2001, *The Doctrine of Double Effect*, Notre Dame: University of Notre Dame Press.

Woollard, F., 2015, *Doing and Allowing Harm*, Oxford: Oxford University Press.

——, 2023, "The New Trolley Problem: Driverless Cars and Deontological Distinctions," *Journal of Applied Philosophy*, 40(1): 49–64.

Zach, N., 2023, *Ethics for Disasters*, London: Rowman and Littlefield.

Zimmerman, M., 2002, "Taking Moral Luck Seriously," *Journal of Philosophy*, 99(11): 553–576.

Academic Tools

- How to cite this entry.
- Preview the PDF version of this entry at the Friends of the SEP Society.
- Look up topics and thinkers related to this entry at the Internet Philosophy Ontology Project (InPhO).
- Enhanced bibliography for this entry at PhilPapers, with links to its database.

Other Internet Resources

[Please contact the author with suggestions.]

Related Entries

consequentialism: rule | doing vs. allowing harm | double effect, doctrine of | ethics: virtue | Kant, Immanuel: moral philosophy | Moore, George Edward: moral philosophy | moral dilemmas

Acknowledgements