A Kantian Approach to Dilemmas: Solving the Trolley Problem

ABSTRACT: In this paper I apply a Kantian insight into the nature of the relationship between law and ethics to resolve moral conflicts such as those in the (in)famous “trolley problem.” Kant argues that in a society where everyone is morally equal, the determinations of legitimate public institutions as to our respective rights and obligations generally take normative priority over private ethical judgments in cases of conflict. This is known as the “priority of right” and its application resolves the original trolley problem posed by Phillipa Foot to rationally vindicate popular intuitions. The Kantian rationale for the priority of right also illuminates the resolution of intractable conflicts between strict obligations (dilemmas). Whether tragic moral conflicts are possible or not, since no one can consent to be subject to the enforcement of contradictory obligations, the prescriptive system of obligations defined in public law must be made consistent. Justice requires dilemmas be resolved.

I. Systematic freedom and the priority of right

A. Systematic freedom

In the Doctrine of Right (DR), Kant defines the "Universal Principle of Right" (UPR) as follows:

Any action is right [just] if it can coexist with the freedom of every other under universal law;

or if on its maxim the freedom of choice of each can coexist with everyone’s freedom in accordance with a universal law (DR: 6:230).

Kant thus defines the legal permissibility (rightfulness) of any action in terms of its consistency with everyone’s equal right of freedom under universal law. If an act can coexist with the equal right of freedom of everyone under a universal law, then it is legally permissible. While Kant defines the legal permissibility of acts here, permissions, duties and (claim-)rights are logically interdefinable by taking any one as a primary operator (see Hohfeld 1919: 35-50).

Kant reiterates that freedom exists in a system under universal law when defining the right of freedom:
Freedom (independence from being constrained by another’s choice), insofar as it can coexist with the freedom of every other in accordance with a universal law, is the only original right belonging to every [person] by virtue of his humanity (DR: 6:237).

Hence while freedom is “independence from being constrained by others,” according to Kant, the right of freedom is that freedom systematically limited by everyone’s equal right of freedom under a universal law. The right of freedom, together with its correlative duties and permissions, thus lacks definition outside a system of equal rights of freedom under universal law.

But how are we to determine the shape and scope of equal rights of freedom in the system? Freedom as independence from constraint is neither self-explanatory nor necessarily self-limiting. Maximizing equal rights of freedom consistently in a system is both vague and ambiguous (Hart 1973: 547-50; see Rawls 1993: 291-92). For example, rules of debate make equal rights to speak consistent by systematically limiting them, but the rules are specified by reference to the goal of a good debate, not by reference to the vague end of maximizing equal freedoms to speak (Hart 1973: 543).1

B. Possible universal consent

Kant's possible agreement gloss on the formula of humanity of the categorical imperative is the key to understanding how justice shapes equal rights of freedom to coexist in a system in accordance with a universal law. According to Kant, public laws that shape duties of right are just if and only if everyone could consent to them (see O'Neill 2011: 170-185). Kant articulates justice as possible universal consent most clearly in his political essay, "On the Common Saying: 'That May Be Correct in Theory but It Is of No Use in Practice’” (T):

---

1 Arthur Ripstein argues that this objection can be met by recasting Kantian freedom normatively as "purposiveness" (see Ripstein, Force and Freedom, (Harvard: Cambridge, 2009), Chapter Two). In my view freedom conceived as purposiveness is deeply flawed, but the issue is outside my scope here.
For this is the touchstone of any public law's conformity with right...if a public law is so constituted that a whole people could not possibly give its consent to it..., it is unjust; but if it is only possible that a people could agree to it; it is a duty to consider the law just... (T: 8:297)

Laws structuring rightful relations with others must be such that it is possible for everyone to consent to them, according to Kant. What this implies is that some positive public law is normatively obligatory to secure necessary conditions for the possibility of universal consent, while the rest of positive public law is permissible in the service of other aims, so long as it does not violate those conditions. Kant's modal standard of legitimacy thus both generates some necessary (obligatory) positive law and simultaneously restricts all possible (permitted) positive law. Necessary law includes constitutional protections for the constituents of freedom, equality, and the rule of law, as well as a number of other constitutional rights and duties.

C. The priority of right

According to Kant, reason alone, however, cannot completely specify a priori what our rights and duties with respect to each other are (DR: 6:312). In a society where everyone is morally equal, each person thus has her "own right to do what seems right and good to [her] and not to be dependent on another's opinion about this," Kant says (DR: 6:312). No one individual has the moral authority to unilaterally define what everyone's moral rights and duties are with respect to others (i.e., legislate them), or to enforce them (i.e., execute them), or to resolve disputes (i.e., adjudicate them).

Hence even if everyone strives to act perfectly ethically with respect to each other, according to Kant, wronging one another in our social interactions is inevitable in the absence of a public authority, since "when rights are in dispute (ius controversum), there would be no judge competent to render a verdict having rightful force" (DR: 6:312). What is required, according to Kant, is to construct
...a system of laws for a people...which because they affect one another, need a rightful condition under a will uniting them, a constitution (constituto), so that they may enjoy what is laid down as right (DR: 6:311, emphasis in original).

Kant refers to this system of public laws and institutions as "public right," and a society existing under such a system as one existing in a "rightful" or "civil" condition, as opposed to one in a "state of nature." Only by constituting a united will to authoritatively determine, enforce, and adjudicate our rights and duties with respect to each other can we avoid wronging one another in cases of conflict, Kant argues.

The determinations of legitimate public authority as to the rights and duties of everyone interacting in community therefore generally take normative priority over individual ethical judgments in cases of conflict. The normative priority is known as the “priority of right.” To reject the authority of public law and institutions and instead use one's own individual judgment in cases of conflict is to act wrongly and unethically, indeed, to commit wrong "in the highest degree," according to Kant (DR:6:308n).

II. Resolving moral conflicts such as in the trolley problem

A. The original trolley “problem:” Driver versus Fat Man

Consider the original ("Driver") version of the "trolley problem" (Foot 1967: 3): Imagine you are driving a trolley whose brakes have failed. The runaway trolley approaches a fork in the tracks, and you must choose which track the trolley will take. On the main track are five people who will be struck and killed if you stay on course, while on the side track is one person who will be struck and killed if you switch tracks. What are you obligated to do? In polls and experiments, most people (about 90%) say they would turn the trolley (see, e.g., Mikhail 2007).
Now contrast Driver with the following variation ("Fat Man") (Thomson 1976: 207-8): Imagine that instead of driving the trolley, you are standing on a footbridge overlooking the tracks. The five are still in the path of the runaway trolley, but now there is no side track. Standing next to you on the footbridge is a fat man leaning over the footbridge railing. You realize that you could stop the trolley and save five people if you push the fat man off the footbridge. He would be struck and killed, but the collision would block the forward momentum of the trolley, saving the five. Should you push the fat man over? Most people (again, about 90%) say they would not do so, in a reverse mirror image of the intuitions in Driver (Mikhail 2007).

The trolley "problem" originally raised by Phillipa Foot is the problem of how to rationally reconcile moral intuitions in Driver with those in cases like Fat Man, since most people are willing to kill one to spare five in the former but not in the latter case (Foot 1967: 3). Foot suggests that "negative" duties such as to avoid injuring or killing others are morally more important than "positive" duties such as to render aid to them (Foot 1967: 4-7). In Driver, Foot says, you are faced with a conflict between negative duties not to kill five and not to kill one, and since you must therefore violate a negative duty not to kill someone no matter what you do, it is better to turn the trolley and kill fewer people (Foot 1967: 5). By contrast, in cases like Fat Man, you are faced with a conflict between a negative duty not to kill one (the fat man) and a positive duty to protect the five from harm. In such cases, the negative duty is more important than the positive one, Foot claims (Foot 1967: 5). One therefore should kill the one to spare the five in Driver but avoid doing so in Fat Man.

*The priority of right solves the original trolley problem*

Foot's analysis is roughly correct but incomplete. To complete the analysis Foot needs to provide some account of why and in what sense "negative" duties to avoid acts such as killing others should take
normative priority over "positive" duties to perform acts such as protecting others from harm (see Thomson 2008: 372). I argue that duties not to kill in Foot’s trolley problem take normative priority not because they are negative duties but because they are duties of right, whereas conflicting positive duties to protect others from harm in cases like Fat Man are ethical duties. Foot's distinction between negative and positive duties roughly tracks the distinction between legal and ethical duties, since most legal duties are negative and most ethical duties are positive duties. But the relevant distinction here is between duties of law and those of ethics.

Perhaps you are one of the 10% who think it might not be unethical for you to push the fat man off the footbridge because that minimizes lives lost. But the fat man's legal right to his life includes at least a right not to be coerced to die in order to aid others. Indeed, this much of his right to life likely must be present in any legitimate system of equal freedom to which everyone could possibly consent (see PP: 8:349-50). Since the shape and scope of the fat man's right to his life has thus already been determined authoritatively in public law, the priority of right entails that you have a normative duty to respect that right, whatever your ethical opinion in the case may be. The moral intuition not to push him is therefore rational.

In Driver, by contrast, there is a conflict between a legal duty not to kill the one and legal duties not to kill each of the five. Some may object that by not turning the trolley, the driver avoids taking action and so avoids violating any duty not to kill the five. But this objection fails because as the driver of the trolley you are subject to a prior legal duty to drive the trolley safely and failing to fulfill this duty constitutes an action by omission. To see this prior duty more clearly, suppose there is no one on the sidetrack. Or compare an analogous case where you are driving a car: if there are five people stranded in the lane ahead and you could safely change lanes to avoid killing them, then choosing to
nevertheless maintain your lane and kill them would violate a prior legal duty to drive the car safely (Thomson, 2008: 369). In cases of conflict between legal duties, the priority of right does not determine the outcome, and this is what distinguishes Driver from Fat Man. Legal duties are violated no matter what one choose to do in Driver; it is therefore intuitively reasonable to apply the fallback principle of minimizing harm in the case. The intuition to turn the trolley in Driver is thus also a rational one.

B. A trolley non-problem: Bystander

Before moving on, it is necessary to distinguish a trolley non-problem that has generated a lot of confusion, a variation referred to as "Bystander." This case is precisely the same as Driver, except that instead of being the driver of the trolley, you are a bystander standing next to a lever that you could pull (or not) in order to turn the trolley to the side track, so killing one and sparing five. Some claim that making this change transforms the moral decision one must make from that between choosing whether to kill one or kill five, as in Driver, into that between choosing to kill one or not to save five, as in Fat Man.

Bystander is a non-problem because it is posed ambiguously. It is in fact unclear in Bystander whether when you choose not to pull the lever to turn the trolley, you are choosing to violate your duty to operate the trolley safely and so to kill five (as in Driver), or choosing merely not to save them (as in Fat Man). In Fat Man, unlike Driver, you have no prior legal duty to prevent the trolley from harming the five, because there is no general legal duty to help or protect others (though of course there is an ethical one). But since the bystander in this thought experiment (bizarrely) exercises a level of control over the trolley’s operation as complete as that of a driver, it is not unreasonable to think that the bystander might be subject to a similar legal duty as a driver to operate the trolley safely with respect to
the five. If the bystander is indeed subject to such a prior legal duty, then the case is like Driver and there is a conflict of legal duties; if the bystander is not subject to such a duty, however, then the case is like Fat Man and the priority of right determines the outcome.

Intuitions in Bystander thus shift according to whether subjects make an analogy with Fat Man or with Driver. In experiments where a case like Fat Man, rather than Driver, is presented to subjects before Bystander, many fewer would still choose to turn the trolley in Bystander, and those who would are much less sure about their decision to do so (see Petrinovich and O'Neill, 1996: 156-8). Such ordering effects in fact confound intuitions in every variation of the trolley problem except Driver and Fat Man (see Liao, et al., 2007). Thought experiments where intuitions shift because duties are posed ambiguously are bad thought experiments.

C. The real trolley problem: Driver

Let us assume, as indeed Foot and Thomson both do, that the conflict in Driver is one between strict duties not to kill each of the five by maintaining course and a strict duty not to kill the one by turning the trolley. Foot argues that it is better to violate only one negative duty not to kill rather than five and that this is why you should turn the trolley in Driver (Foot 1967: 5). But while minimizing harm may be a reasonable ethical principle, the principle is contestable as a legal one. Principles of justice characteristically bar the violation of one person’s rights to achieve a greater good such as to save many people; hence it is not clear why justice should allow the violation of one person's rights to achieve the greater good of avoiding violating five people's rights. The one whose rights are violated might

---

2 The graphic on this page in the cited source unfortunately displays the relevant result erroneously, as is evident from a reading of the text itself, which is clear. An experiment isolating and replicating this result would be useful.
complain of being wronged in either case. The conflict in Driver therefore appears to be a genuine dilemma of conflicting strict legal obligations.

**Implications of justice for dilemmas like Driver**

I offer the following approach to the resolution of the dilemma in Driver. First, I stipulate that the problem is indeed a dilemma in which the driver is subject to contradictory strict legal obligations not to kill the one and not to kill each of the five. That is, there is no other legally relevant factor, such as the act-omission distinction, or a superior right on one side or the other due to fault, that would eliminate or prioritize one of the obligations. I then appeal to Kant's normative requirement that public laws are those to which everyone could possibly consent. What does the normative requirement of possible universal consent imply in a case like Driver?

The first implication is that neither legal obligation in the dilemma can be rightfully enforced. It is not possible to consent to be subject to the enforcement of contradictory strict legal obligations, as this is tantamount to consenting to arbitrary acts of coercion. The second implication is that the dilemma must be resolved by public authority, either by legislative action or by judicial or executive order. It does not matter how the conflict is resolved, so long as the procedural and substantive requirements of justice are met when resolving it. What matters is that the conflict is resolved; and moreover, its resolution may legitimately vary by jurisdiction, so long as requirements constitutional of justice are met. Variation in the law by jurisdiction is a common feature of most legal systems: in some U.S. states, for example, contributory negligence completely bars recovery by injured plaintiffs, while in other states, fault might play no or a very limited role. Yet in each state, the law that resolves the legal conflict is rightfully enforceable because it violates no constitutional law.
From the point of view of justice, then, dilemmas such as in Driver are little different from other conflicts of rights in the system of equal freedom. Just as in other conflicts, the determinations of public authority as to their resolution will take normative priority over varying private ethical judgments. The main difference is that in dilemma cases we assume that there is no clear rational resolution of the conflict at issue, whereas in ordinary cases of conflict, we assume that some rational resolution of the conflict may exist. Regardless, the rationale that requires their resolution remains the same: without a public authority to determine our respective obligations in social interactions, we inevitably wrong one another. Hence when public authority does resolve such conflicts, we have a moral duty to respect that resolution, even if we disagree with it.
References


Kant, I. (1992) In P. Guyer and A. Wood (Eds.), *The Cambridge Edition of the Works of Immanuel Kant*. Cambridge: Cambridge University Press. All references to Kant's work are from the Cambridge edition unless otherwise noted. Citations are according to standard Academy pagination.


