Research Square

Preprints are preliminary reports that have not undergone peer review.
They should not be considered conclusive, used to inform clinical practice,
or referenced by the media as validated information.

Robust Deontology Without Paradox

Alec Walen

awalen@rutgers.edu

Rutgers, The State University of New Jersey

Daniel Wellisch

Research Article

Keywords:

Posted Date: November 17th, 2026

DOI: https://doi.org/10.21203/rs.3.rs-11598278/v1

License: This work is licensed under a Creative Commons Attribution 4.0 International License.
Read Full License

Additional Declarations: No competing interests reported.

Robust Deontology Without Paradox

Lessons on the Structure of Deontological Valuing[*]

Richard Chappell offers an argument purporting to establish a new paradox for deontology.[1] The new paradox he claims to have found concerns a supposed contradiction between the preference agents should have not to wrong others themselves and robustness.[2] Robustness, as he uses the term, is a quality of deontic constraints: that agents should not only prefer to respect deontological constraints themselves but should also prefer that others respect those same constraints. Chappell thinks he can show that deontologists must take deontological constraints to apply only to themselves or recognize that "deontology is false" (178).

We provide a straightforward criticism of Chappell's argument, showing that, as he articulates it, it is invalid because it relies on an unjustifiable claim about transitivity. We accept that the cases he relies on fall into the preference hierarchy that he assigns to them, but his argument requires both a quantitative and a qualitative dimension. We call that strong transitivity, and his argument fails at that level.

But we do not content ourselves with pointing out this simple failure, which is independent of the nature of deontology. We want to draw attention to two substantive misconceptions about deontology that affect Chappell's argument. We think these are of general interest to anyone looking to properly understand deontology, including those who wish to weigh its appeal against consequentialist alternatives.

---

[*] Thanks to _____ for their comments on an earlier draft.

[1] Richard Y Chappell (2025). Preference and Prevention: A New Paradox of Deontology. *Free & Equal: A Journal of Ethics and Public Affairs*, 1: 175–201. DOI: https://doi.org/10.16995/fe.18062. Page references in parentheses refer to this article.

[2] The original paradox of deontology was framed by Robert Nozick, who asked the question: "If nonviolation of C is so important, shouldn't that be the goal? How can a concern for the nonviolation of C lead to the refusal to violate C even when this would prevent other more extensive violations of C?" *Anarchy, State, and Utopia*. Basic Books (1974), 30. This is easily answered by noting that constraints standardly reflect rights, and rights are relational. For example, if A aims to use five people as a means of achieving some aim, killing them in the process, and B could save those five by using C as a means of stopping A, killing C in the process, there is nothing paradoxical in saying that B may not kill C. The five have claims on B to be saved from A; C has a claim on B not to be lethally used as a means. There is no reason why the latter claim cannot be stronger than the collection of former claims, both *on B*. The mistake in the original paradox was appealing to an agent-neutral notion of value or "importance," rather than looking at the claims that compete on each agent. See _____. For a list of three other supposed paradoxes with deontology, see "Preference and Prevention", p. 177, n. 4.

The first misconception concerns the strength of preferences for respecting deontological constraints, as compared to violating them for the sake of consequentialist benefits. Chappell's interpretation of robust deontic preferences assumes that this preference can be quite weak. He makes this assumption in an effort to capture all deontologists who care not only about what they do but about what others do. But in truth, his interpretation is too weak to do justice to core deontological commitments. Curiously, if we strengthen deontological preferences to do justice to core deontological commitments, the problems with strong transitivity go away. But that does not help his argument because of his second misconception about deontology.

The second misconception is that it makes sense to talk about a "generic murder". Chappell's argument trades on the value to a bystander of a "generic murder" being committed. But for a deontologist, there is no such thing as a "generic murder". Deontologists should be ready to accept that one wrongful killing of a stranger is just as bad as another, when they stand in the same relationship to the actor. But the disvalue of a murder to a person assessing the value of the actions of another depends crucially on where it occurs in relationship to the agent whose actions are being assessed. Deontology is primarily concerned with the claims of patients, but these are claims *on* agents. To avoid confusion, we think it is helpful to describe deontology as substantially patient-claim-*grounded*, but agent-*centered*.[3] This is as true for assessing the actions of others as for assessing one's own actions. Overlooking this fundamental structural point about deontology is the fundamental substantive misconception about deontology that makes Chappell's argument substantively unsound.

We proceed as follows. In section I, we lay out Chappell's argument for a new paradox of deontology. In section II, we critique it in its own terms. In section III, we reconstruct the argument with stronger preferences that match better substantive deontological commitments. Finally, in section IV we explain why the reconstructed argument does not help Chappell because it still relies on the misguided notion that a deontologist should accept the idea of an extra "generic murder" in the world.

I. Chappell's Argument

In this section, we start with a brief recap of Chappell's defense of robustness, with which we agree. We then describe the set of cases he relies on and his reductio, summarizing his defense of it.

[3] In defense of this conception of deontology, see _____'s work on the _____. See _____, _____.

a. In Defense of Robustness

A robust concern that each agent should have for whether other agents violate the rights of yet other victims is attractive for a number of reasons. As Chappell points out, it "preempts any possible charges of egoism, self-indulgence, or 'clean hands' fetishism, fitting well with an attractively 'patient-centered' [patient-claim-grounded] conception of moral concern" (179–180). As he also points out, "[i]f you take deontic constraints seriously, you should not want the agent" who is in a position to violate the rights of others to do so, not even if doing so is the only way to prevent more rights violations (180). This might seem a paradoxical preference to a utilitarian or agent-neutral consequentialist,[4] but it is a central tenet of standard deontological views.

Furthermore, if one cared only about one's own wrongdoing and not others', one would regret that another agent does not perform the action that wrongs another but minimizes wrongdoing overall. But one should not wish others to engage in wrongdoing. As Chappell puts it, "only the robust view respects the datum that moral perfection is not lamentable" (181).

b. Cases and Chappell's Reductio

Chappell relies on the following four cases to set up his reductio (189). All are based on the background assumption that an agent that he calls Protagonist confronts a situation in which there are "five other agents who are each about to murder a different innocent victim" (189), and Protagonist thinks she may be able to save the five by killing a sixth individual.

> **Five Killings**: Protagonist does nothing, so the five other murders proceed as expected.
>
> **One Killing to Prevent Five**: Protagonist kills one as a means, thereby preventing the five other murders.
>
> **Failed Prevention**: As above, Protagonist kills one as an intended means, but in this case fails to achieve her end of preventing the five other murders. So all six victims are killed.

[4] Indeed, this thought is the core of Nozick's original "paradox" of deontology. See note 2.

Six Killings: Instead of attempting to save the five, Protagonist simply murders her victim for the sheer hell of it, just like the other five murderers. So all six victims are killed.

Deontologists believe that they should not kill one to prevent five murders, so if they are to hold a robust view, they must also prefer Protagonist not to kill one as a means of preventing five murders. Therefore, they should prefer Five Killings over One Killing to Prevent Five. Clearly, they should also care about the lives of the five. So *if* Protagonist decides to act wrongly and kill the one, they should at least hope that this prevents the five murders. That is, they should prefer a world in which, if Protagonist wrongs one to save the five, she has not done so in vain. That implies that they should prefer One Killing to Prevent Five to Failed Prevention. And the importance of agents not being motivated badly should lead them to prefer Failed Prevention to Six Killings.

This brings us to Chappell's formal argument. It builds on this preference hierarchy and adds a quantitative dimension, marked by the idea of "vastly" preferring. Importantly, it is all framed from the point of view not of Protagonist but of "a fully informed, morally ideal, neutral bystander" (186) (from hereon, "Bystander").

To formulate the argument, he introduces four symbols or symbol sets:

- $X \succ Y$ means that X is preferable to Y
- $X \succcurlyeq Y$ means that X is preferable to Y or is equivalent to Y
- $X \succ\!\!\succ Y$ means that X is "vastly" preferable to Y, where that means that the preference is "strictly stronger than" (190) the preference that Bystander should have for "one less generic murder to occur in the world" (190)
- $X \lozenge\!\succ Y$ means that it is permissible (though not required) to prefer X to Y.

With that, here is his argument (190):

1) Protagonist acts wrongly in One Killing to Prevent Five, due to violating an important deontic constraint, and ought instead to bring about the world of Five Killings. (For reductio)
2) If an agent can bring about just W1 or W2,[5] and it would be wrong for them to bring about W1 (but not W2) due to violating an important deontic constraint, then $W2 \lozenge\!\succ W1$. (Weak robust constraints)
3) Five Killings $\lozenge\!\succ$ One Killing to Prevent Five. (From 1, 2)
4) One Killing to Prevent Five $\succ\!\!\succ$ Failed Prevention. (Premise)
5) Failed Prevention $\succcurlyeq$ Six Killings. (Premise)

[5] W represents a possible world.

6) Five Killings ◊≻≻ Six Killings. (3–5, transitivity)
7) It is not the case that Five Killings ◊≻≻ Six Killings. (Definition of '≻≻')
Contradiction (6, 7)

We will now give a compressed defense of this argument. We jump directly to line (3), which represents a weak version of a robust deontological commitment, namely that it is at least permissible for Bystander to prefer that Protagonist not kill one even if killing one is the only means to save five from being murdered.[6]

The justification for line (4) is, as Chappell puts it: "on any sane view, One Killing to Prevent Five must be vastly preferable to Failed Prevention. In the former scenario, Protagonist kills one person to prevent five other killings. The latter scenario contains this killing, and everything that's morally objectionable about it, plus five additional, completely gratuitous murders" (191). One Killing to Prevent Five must be *vastly* preferable to Failed Prevention because of the definition of "vastly superior": X is "vastly" preferable to Y "means that the preference is strictly stronger than the preference that Bystander should have for one less generic murder to occur in the world."[7] Failed prevention involves *five* more generic murders occurring in the world—five in addition to the one wrongful killing Protagonist would commit in either case—so the preference should be strictly stronger than for having merely one more generic murder.

Line (5) captures the thought that motives might matter to preferability, and if they do, then Failed Prevention is preferable to a world in which the agent is "evilly motivated" (192).

Line (6) (supposedly) follows by transitivity.

Line (7) (supposedly) follows from the definition of ≻≻ (i.e., one option being vastly preferable to another). Chappell takes line (7) to be true because he thinks (a) that the number of generic murders in the world of Five Killings is exactly one less than the number of generic murders in the world of Six Killings, and (b) that for "Five Killings ◊≻≻ Six Killings" to be true, there must be more that would make Five Killings preferable to Six Killings than simply one less murder. Given these assumptions, he can validly conclude that it is not the case that Five Killings ◊≻≻ Six Killings.[8]

---

[6] We will argue in the next section that a deontologist should use the ≻≻ relationship here, not the weaker ◊≻ relationship.
[7] From the list of definitions above, internal quotation marks and page citations omitted.
[8] One might be tempted to object that line (7) is false on the ground that Protagonist has a reason to strongly prefer Five Killings to Six Killings. It's not merely the case that there is one fewer killing in Five Killings; *she* is not behaving as a murderer. That should be very weighty with her. But this objection loses sight of the fact that

Finally, the contradiction between lines (6) and (7) is obvious.

* * *

Chappell draws the following conclusion from the contradiction he has derived (192):

> The challenge for deontologists is that there just isn't enough moral room between Five Killings and Six [Killings] to accommodate the moral gulf that ought to lie between One Killing to Prevent Five and Failed Prevention. As a result, when they allow Five Killings to be morally preferred to One Killing to Prevent Five, they are unable to accommodate our moral datum (4), that Failed Prevention is vastly dispreferable to One Killing to Prevent Five.

We agree to this extent: if strong transitivity works to produce line (6) and line (7) is true, then the deontologist faces a contradiction, and it makes sense that they should give up on line (3), i.e., the possibility of robustness, or line (1), the premise that it is wrong for Protagonist to kill one to save five. Fortunately for deontologists, the argument for line (6) involves an invalid inference and line (7) wrongly presupposes that deontologists can talk about a "generic murder". To defend these claims, we start by exposing the problematic use of strong transitivity to derive line (6).

II. Rejecting the Reductio

The problem with deriving line (6) via transitivity is that strong transitivity does not work when there is no common metric, and lines (3) through (5) each involve a different metric. To name them:

- Line (3) compares the value of two different kinds of moral concerns: respecting deontological constraints and outcomes;
- Line (4) compares the value of one kind of moral concern: outcomes; it also provides a measure of vast preferability; and
- Line (5) compares the value of yet another kind of moral concern: subjective wrongdoing.

Running a more-preferable-than relationship across all three kinds of concerns, we find that the vastly preferable relationship in line (4) drops out. To demonstrate that point, we will present an analogy in which preferability is cashed out in terms of value to a

---

the preferability judgment in this argument, which is designed to test the idea of robustness, is made not from Protagonist's point of view but from Bystander's point of view.

department at a research university that is considering whether to hire a job candidate. We will spell it out with analogous line numbers:

3*) Superstar scholar but poor teacher $\succ$ poor scholar but a great teacher. (Premise supported by the choices of many research universities' hiring or promotion committees.)

- Here we have two different kinds of considerations traded off against each other, as in the original line (3). It is a stronger relation, as we have dropped the ◊ qualifier. But that should not be a problem in the context of testing transitivity.

4*) Poor scholar but great teacher $\succ\succ$ poor scholar and poor teacher. (Premise, which is obviously true with regard to mere preferability. We assume that we can use, for the sake of argument, some intuitive reasonable metric to establish that great versus poor teaching grounds a "vastly preferable" relationship that applies in the context of this argument.)

- Here we have a single consideration in play, just as in the original line (4).

5*) Good at committee work $\succ$ bad at committee work. (Premise, which is obviously true, as long as all else is equal. We assume that committee work is less significant than scholarship and teaching, so the good versus bad quality grounds only a preferable, not a "vastly preferable" relationship.)

- Here we have a new kind of consideration, as in the original line (5). Once again, we have dropped a weakening in the original argument, which used the $\succcurlyeq$ in line (5). But once again, this should not be a problem if we are just testing transitivity.

Now, the question is: Can we infer that a candidate who is a superstar scholar but a poor teacher is vastly preferable to one who is bad at committee work? We can't without knowing how good at scholarship and teaching the one who is bad at committee work is. But if we want to track Chappell's original argument, we should add that information. That's because we read Chappell as tracking a single character, Protagonist, through the various features she or her actions might have, changing only what is explicitly described as different. If we do the same with the academic candidate, then we should validly infer that a candidate who is a superstar scholar and poor teacher who is good at committee work is at the very least preferable to the same candidate if she were a poor scholar and poor teacher who is bad at committee work.

The important question, however, is: Can we infer that the candidate who is a superstar scholar and poor teacher who is good at committee work is *vastly* preferable to the same candidate if she were a poor scholar and poor teacher who is bad at committee work? If we can, then we would be able to infer an analogous line (6*), though one slightly stronger than the original line (6) because it would be lacking the permissibility ◊ qualification. Transitivity, as Chappell uses it, suggests that we can derive an analogous line (6*) for this academic case. But if we look more closely, we see a problem.

Note that in the academic analogy, we defined the ≻≻ relation in terms of the value of great teaching versus poor teaching, all else equal. But when we compare the first version of the candidate to the last version of the candidate, that difference is gone. Both versions are poor teachers. The differences are that the first version is an excellent scholar and the last is a poor scholar, and that the first is good at committee work while the last is poor at committee work. What can we infer from that? We know that great scholarship outweighs poor teaching by the ≻ relation. But we haven't got a definition of just how strong that relationship is. We also know that being good at committee work outweighs being bad at committee work by the ≻ relation. But we don't know how those two relationships combine. Maybe they add up to the ≻≻ relation. But nothing about the premises allows us to *infer* that as a matter of logic. Therefore, we cannot infer an analogous, strong version of line (6*).

Can we infer the weaker version of candidate comparison?

6*) Superstar scholar, poor teacher, and good at committee work ◊≻≻ poor scholar, poor teacher, and poor at committee work

It is intuitively plausible that we can. Recall that X ◊≻≻ Y means only that it is permissible to vastly prefer X to Y. The uncertainty in the way the preference relationships transmit across the cases might seem to imply that it is *permissible* to vastly prefer a candidate who is a superstar scholar, poor teacher, and good at committee work to a version of the same candidate who is a poor scholar, poor teacher, and poor at committee work. But it would be a mistake to move from uncertainty to permissibility. They are different notions. If we don't *know*, given the premises we have accepted so far, that X ≻≻ Y, it may be that further investigation would show that it is not. In that case, it would not be permissible to take X to be vastly preferable to Y. So, we cannot derive the equivalent of line (6).

To see why the analogy works, recall that the ≻≻ relation is defined by Chappell in terms of outcomes. Specifically, the preference for one state must be strictly stronger than the preference to avoid a particular outcome: one extra generic murder. But as we move from step to step, the outcome-to-outcome relationship of line (4) drops out. What we

know is that respecting constraints $\lozenge\succ$ extra murders, from line (3), and that acting on good reasons $\succcurlyeq$ acting on bad reasons, from line (5). We don't know just how much respecting a constraint is preferable to violating it. Protagonist on the good side of line (3) obeys a constraint (in analogy to being a superstar scholar), and Protagonist on the bad side of line (5) flouts the constraint (in analogy to being a poor scholar) *and* acts on bad reasons (in analogy to being bad at committee work). Does logic imply that those two differences must add up to the $\succ\succ$ relationship? No more so than in the academic analogy.

In other words, we see no reason for *strong* transitivity to work when moving across different kinds of considerations like this. Transitivity seems to work, but not *strong* transitivity.

Chappell might reply that strengths of preference can be compared across any pair of options, so that if $X \succcurlyeq Y$ and $Y \succ\succ Z$, then $X \succ\succ Z$ by monotonicity alone. But that reply assumes what is at issue: that the strength of preferences is, or logically should be, independent of the grounds for the preferences. We see no reason to accept that. When the ground for judging one thing vastly preferable to another drops out, we see no reason to think that it must be preserved simply because it is possible to construct a chain of preferences that used it. Thus we reject the assumption that strong transitivity can be relied on with supporting argument that it applies in the case at hand. We think it does not apply in Chappell's argument. And without strong transitivity, Chappell cannot derive line (6).[9]

This is not a surprising conclusion. As Larry Temkin has argued, when "there is no common linear scale on which we can rank, and hence compare, alternatives with respect to an underlying one-place property Q … [this] opens up the possibility that S is a nontransitive relation".[10] This is true of notions like "better than" which can be articulated in terms of a variety of underlying considerations. It can likewise apply to preferability based on several different relations. To repeat, we are not denying that preferability is in fact transitive in the two cases we have been discussing: Chappell's cases involving killing or our analogy involving departmental hiring. Indeed, when we first introduced Chappell's four cases, we accepted his preference hierarchy between them. What we are denying is that the quantitative degree of preferability is transitive. Given that the kinds of comparisons made are different, and that there is no common metric for assessing degrees of preferability, that should not be surprising.

[9] Chappell argues in favor of transitivity of preferences (193–194). But he doesn't engage the question whether *strong* transitivity works.

[10] Larry Temkin (2012). *Rethinking the Good.* Oxford University Press, 170.

III. Vast Preferability Can Be Recovered with a Better Understanding of Deontology

One way Chappell could respond is by strengthening the relationship in line (3). The original argument says:

3) Five Killings $\lozenge\succ$ One Killing to Prevent Five.

But Bystander, as a deontologist concerned with how Protagonist acts, would surely want a stronger relationship than that.[11] Indeed, Bystander would presumably want to argue that the relationship has to be at least the $\succ\!\!\succ$ relationship. And if it *is* the $\succ\!\!\succ$ relationship, then we can recover line (6); indeed, we can derive a stronger version of line (6) without the permissible preference qualification.

Why should it be at least the $\succ\!\!\succ$ relationship? The thought is that Bystander, as a deontologist, would prefer Protagonist not to kill another as a means, and violate his fundamental right not to be so used, not only if five were at risk of being murdered, but even if *many more* were at risk of being murdered.[12] Since this preference is strictly stronger than the addition of more than one generic murder, we should be willing to adopt this much stronger version of line (3):

3**) Five Killings $\succ\!\!\succ$ One Killing to Prevent Five

Looking ahead in our argument to the criticism of the idea of generic murders, it may seem odd that we embrace this argument for line (3**). But we are still here working within Chappell's framework, and within that framework, this is a perfectly reasonable thing to say. Indeed, obeying deontological constraints is normally taken by deontologists to be preferable to preventing any number of murders that others might commit,[13] as long as that number is below the threshold for threshold deontology. Most deontologists today are threshold deontologists, but there is good reason for them to insist that the ratio be *much* higher than five to one.[14] Thus we think that line (3**) better reflects deontological commitments than Chappell's original line (3).

---

[11] Chappell weakened line (3) to mere permissible preference to try to capture the weakest form of robust deontology in his net. See his gloss on line (2) in the blog post announcing the article: the argument "still goes through if you so much as permit any neutral party to prefer the 'deontological' outcome." Richard Y. Chappell, "The Curse of Deontology," Good Thoughts (June 2025), n. 7, https://www.goodthoughts.blog/p/the-curse-of-deontology.

[12] We explore this reasoning further in the next section.

[13] The qualification regarding others committing the murder is crucial, as we point out in section IV.

[14] On the nature and difficulties of threshold deontology, see Larry Alexander (2000). Deontology at the Threshold. *San Diego Law Review, 37*: 893–912. On the thought that the ratio should be much higher than five to one, see _____.

This vastly preferable relationship in line (3**) should be recognized both by Protagonist herself and by any deontologically minded observer who has a robust view of deontology. And with this new premise, we can recover a strong version of line (6). The argument then looks like this:

3**) Five Killings $\succ\succ$ One Killing to Prevent Five

4) One Killing to Prevent Five $\succ\succ$ Failed Prevention

5) Failed Prevention $\succcurlyeq$ Six Killings

6**) Five Killings $\succ\succ$ Six Killings (3**–5, transitivity).

Here's the basic logic behind that deduction: In Five Killings, Protagonist respects a deontological constraint that is vastly stronger than the pull of saving any sub-threshold number of people from murder by others. If Protagonist were to kill one to prevent five killings, she would do something *vastly* dispreferable. Suppose she does that; it would still be vastly preferable if she at least prevents five killings, rather than failing to save them, as in Failed Prevention. And if she were to add malice to her killing, it would be worse still than Failed Prevention. Crucially, we don't have to worry that the vastly preferable relationship in line (4), which concerns outcomes with no change in Protagonist or her action, is based on a dimension of comparison that is absent in line (6**). The vastly preferable relationship appears in line (3**), which anchors the quantitative assessment in line (6**).

To confirm this, consider again the academic analogy. Imagine that the department gave vast priority to good scholarship over good teaching. Unless we reach the threshold at which the candidate can't teach at all, a good scholar is vastly preferable to a bad scholar, no matter how poor his teaching. If we then compare a candidate who is a good scholar with the same candidate turned into a bad scholar who is also a bad teacher, we can be sure that the department would consider the good scholar version vastly preferable to the bad scholar version. If we add that the bad version of the candidate is bad at committee work, we should have no doubt about how strong the preference for the good scholar would be: it would be vastly strong.

IV. Exposing the "Generic Murder" Mistake

Even if we can reconstruct a strong version of line (6) by providing a strong version of line (3), Chappell's argument does not work because line (7) is false.

7) It is not the case that Five Killings $\lozenge\succ\succ$ Six Killings. (Definition of '$\succ\succ$')

Chappell claims that line (7) is true because the difference between Five Killings and Six Killings is only that there is "one less generic murder [that] occur[s] in the world" in Five Killings than in Six Killings (from the definition of "$\succ\succ$"). But Chappell thinks he can say that only because he is relying on a notion of a "generic murder", a notion that ignores a key structural difference between two ways a murder might appear in a scenario.

To bring out what we mean by a structural difference, we introduce "deontological force diagrams" that make it clear how different scenarious should be understood and compared. To make the diagrams easier to read, we start with a simple component to all of the diagrams, a diagram representing the claims of the five victims on Protagonist (Figure 1), along with a general key for interpreting the diagrams (Figure 2):

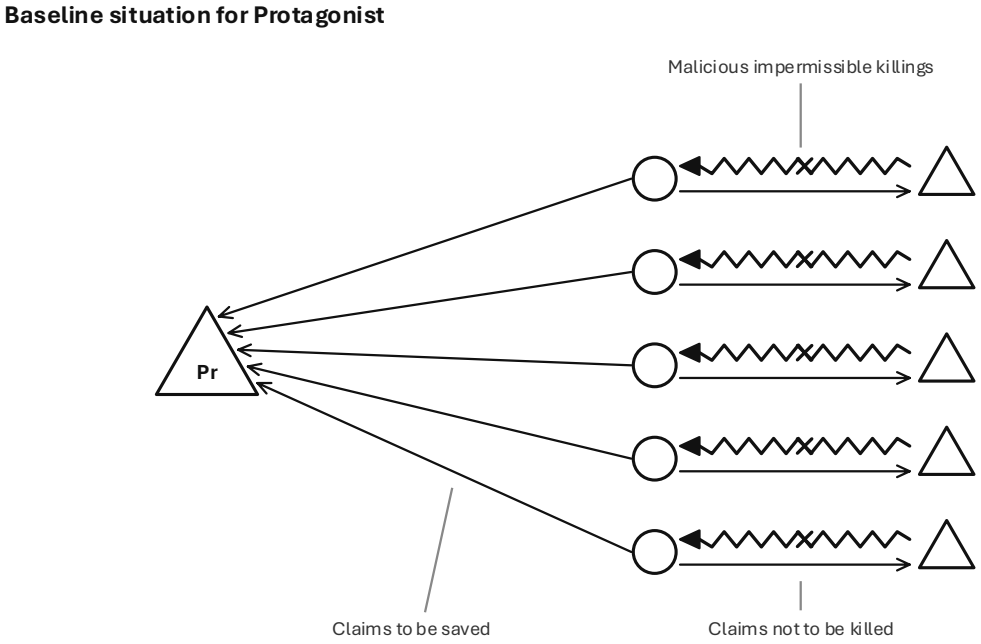
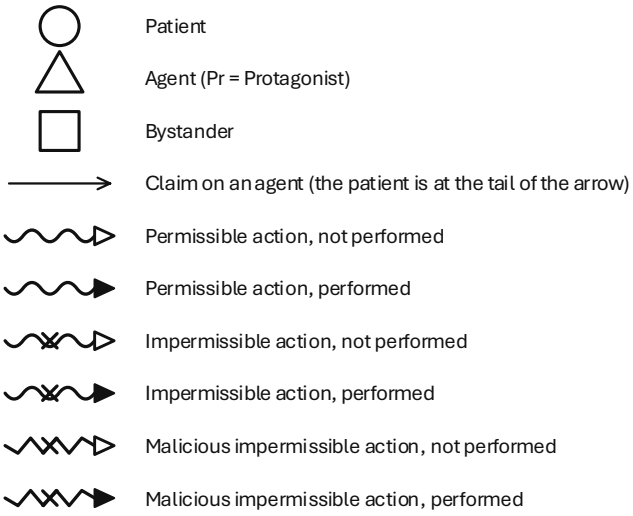


Figure 1. The baseline situation for Protagonist: each of five victims has a claim on Protagonist to be saved and a claim on a bad agent not to be killed.

Key to the deontological force diagrams



Relations between cases, from Bystander's standpoint

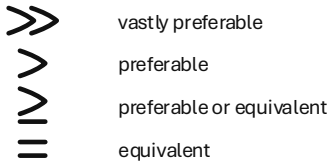


Figure 2. Key to the deontological force diagrams.[15]

[15] For the sake of completeness, we include in the Key signs for permissible actions, though, as a matter of fact, none of the cases involve any permissible actions.

With these basics, we can represent the first comparison, between Five Killings and One Killing to Prevent Five as follows:

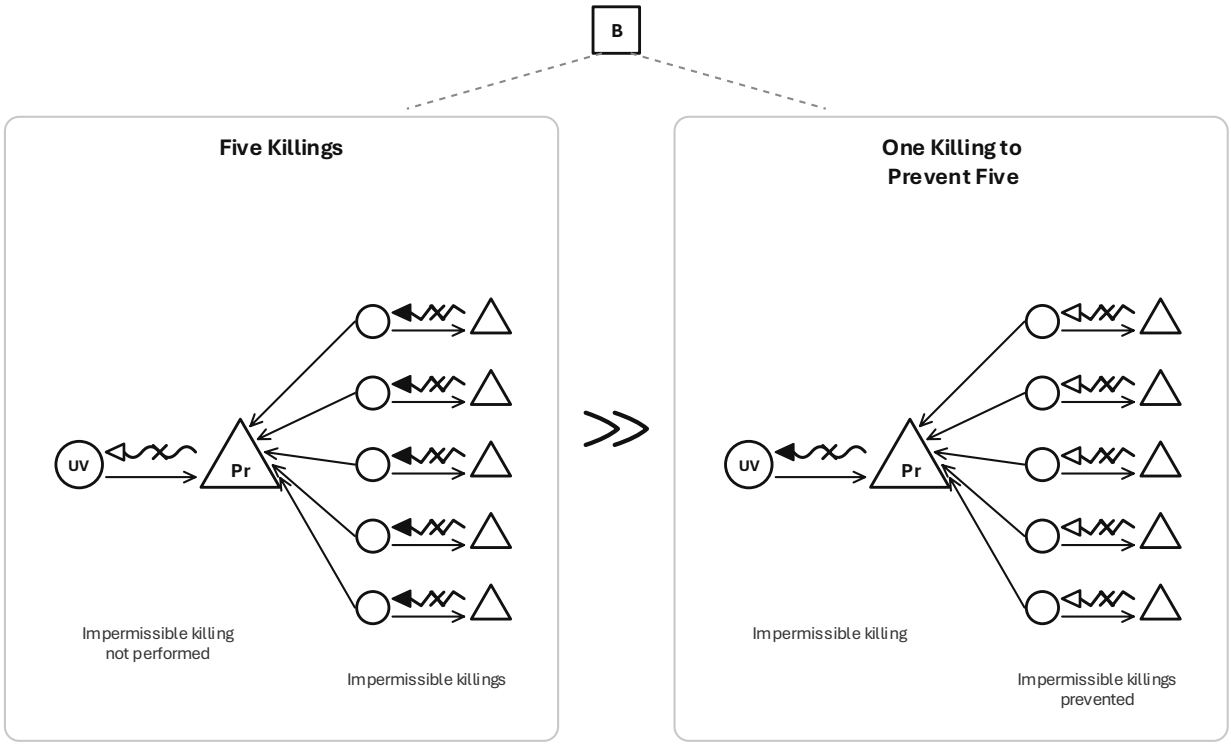


*Figure 3. Five Killings compared with One Killing to Prevent Five, as in line (3**).*

Note, UV stands for (potentially) Useful Victim. In Five Killings UV is not killed, and the five other victims are therefore not saved. In One Killing to Prevent Five, UV is usefully killed, thereby saving the five other victims. And in Failed Prevention (not diagrammed), UV turns out not to be useful, though Protagonist acted on the assumption or hope that his death would be useful.

Now, one way to understand why Five Killings ≻≻ One Killing to Prevent Five is to compare Five Killings with a case in which two more malicious killings are added to the right side of that diagram (as discussed in the previous section; see Figure 4):

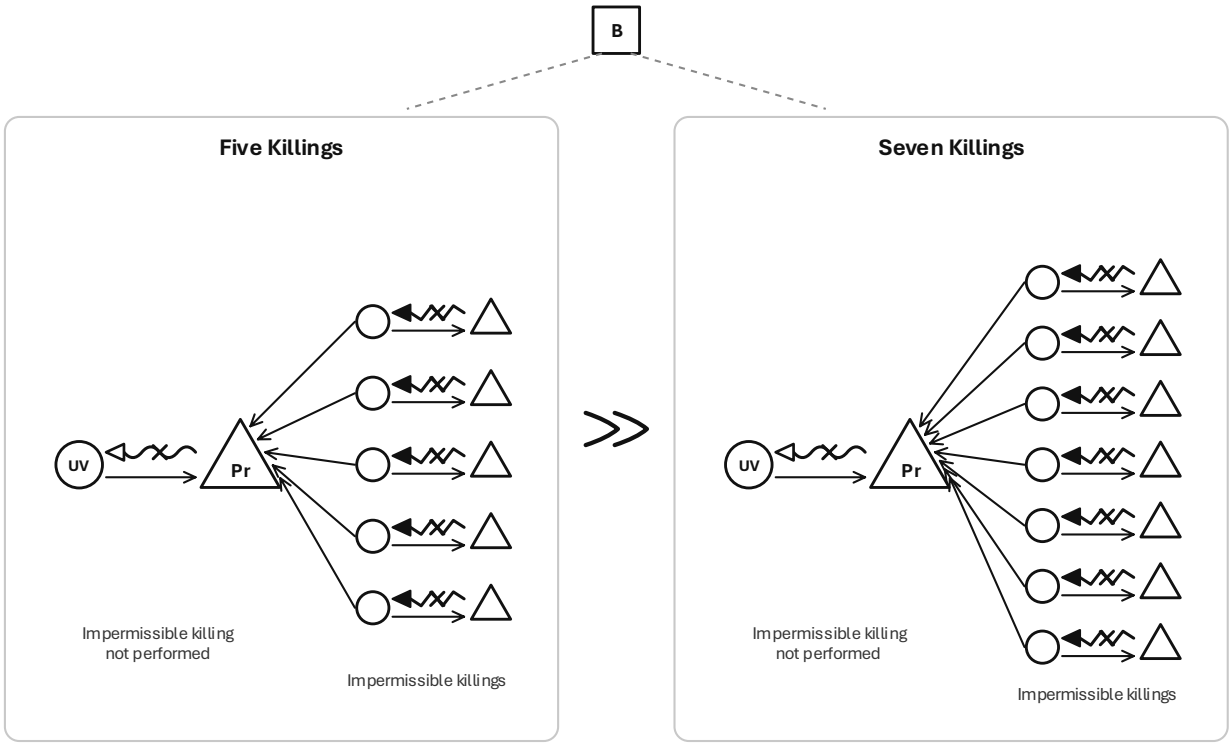


Figure 4. Five Killings compared with Seven Killings.

Five Killings is vastly preferable to Seven Killings in exactly the terms Chappell lays out: Seven Killings involves two extra "generic murders", and the preference for avoiding that should be strictly stronger than the preference for "one less generic murder to occur in the world" (190). To be clear, we do not mean to endorse the concept of a "generic murder" here. But we use the label to track the way Chappell used it in comparing One Killing to Prevent Five and Failed Prevention. That is, we use it to track the results of Protagonist's actions on the killings that *others* perform.

There are three things to learn from this comparison. First, we could produce another set of diagrams like those in figure 3, only these diagrams would depict Seven Killings and One Killing to Prevent Seven. Assuming that we are still far from the threshold for threshold deontology, Seven Killings would still be preferable, if not vastly preferable, to the alternative in which Protagonist kills one to save the seven.

Second, this comparison reinforces the judgment in figure 3, that Five Killings is vastly preferable to One Killing to Prevent Five. The logic is as follows:[16]

i. Five Killings $\succ\succ$ Seven Killings
   - (As just established)

[16] This discussion redeems the promise in note 12.

ii. Seven Killings ≻ One Killing to Prevent Seven (entailed by the nature of deontological constraints un
    - (This is entailed by the nature of deontological constraints under the threshold for threshold deontology; the relation is at least this strong, if not ≻≻)

iii. One Killing to Prevent Seven ≻≻ One Killing to Prevent Five
    - (It might seem that the relation here is one of equality, as both involve simply a single wrongful killing, but we think one should look at how much Bystander should prefer to prevent one of these wrongful actions if he could only prevent one; given the consequences, he should vastly prefer to prevent One Killing to Prevent Five as that would at least result in two more murders being prevented).

From this, it follows by transitivity that

3**) Five Killings ≻≻ One Killing to Prevent Five.

This is a valid use of strong transitivity, unlike in Chappell's initial deduction, because the cross-kind valuation in this line (ii) is not doing the heavy lifting to get us ≻≻ in line (3**). We know from the nature of deontological constraints that Five Killings ≻ One Killing to Prevent Five, and we know from basic transitivity the the preference is stronger than in line (ii). Moreover, it is stronger *in virtue of* these vastly stronger preferences, making it a vastly stronger preference. The difference between this deduction and Chappell's is that in Chappell's initial deduction the kind of relationship that involved a vastly stronger preference was not retained in the conclusion (line (6)); here the comparison implies that the preference relationship is stronger than that in a vastly stronger preference relationship, so the quantitative relationship is retained.

The third lesson is the most important, however. It is that there are structurally different places to add extra killings. Where the extra killings were added in Seven Killings, as compared to Five Killings, is *not* the only place where extra killings could be added. To see that, consider a new case, where Protagonist herself kills someone *else*, the Alternate Victim (AV), not UV. We call this case Five Killings Plus One. We claim that merely by adding one extra malicious killing into the picture at the hands of Protagonist, rather than at the hands of some other agents, matters. Merely adding one extra killing by Protagonist makes her action *vastly* inferior. This is in contrast with adding just *one* extra killing that could be prevented by Protagonist killing UV. That case, which we call One Extra *Other* Killing, would *not* be vastly dispreferable to Five Killings, as it would involve only *one* extra murder that is not prevented. To appreciate the difference, compare the Diagram for One Extra *Other*

Killing (not presented, but modeled on Seven Killings) with the following set of Diagrams comparing Six Killings and Five Killings Plus One (Figure 5):

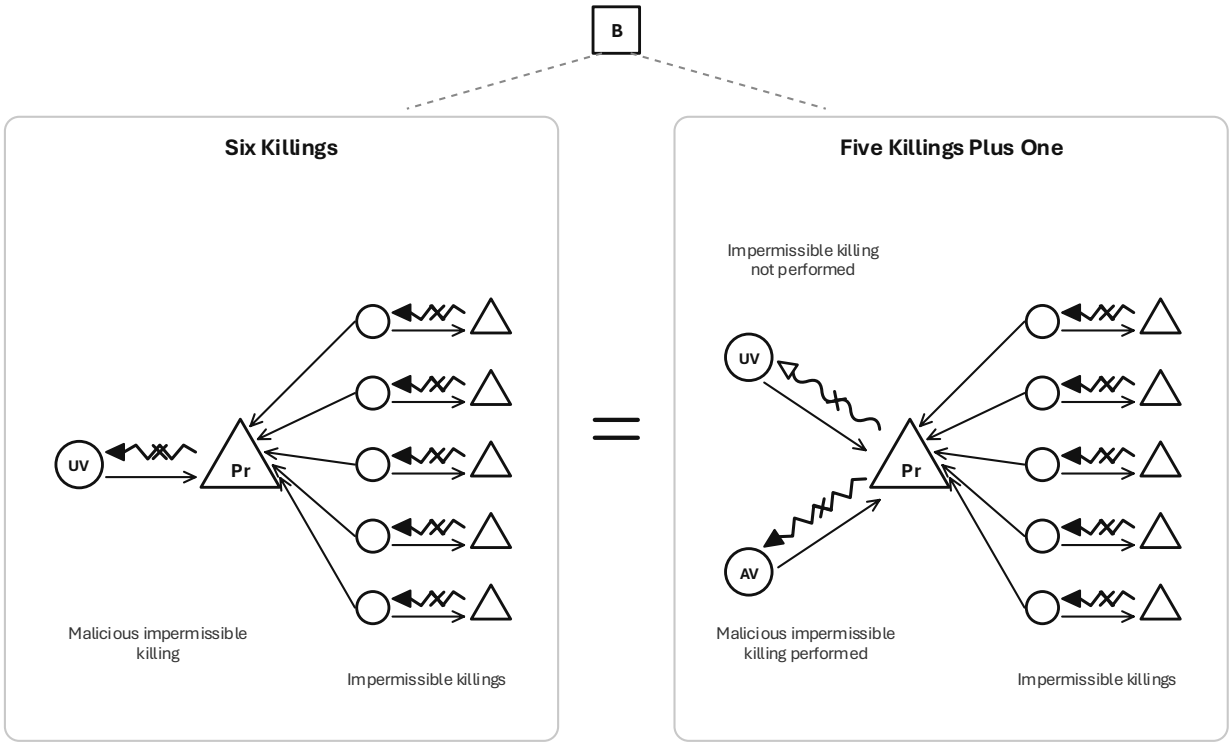


Figure 5. Six Killings compared with Five Killings Plus One.

Bystander has no reason to prefer (to want to prevent if given the opportunity) the killing of UV (when it's in fact useless and maliciously intended) to the killing of AV. UV and AV are equally victims of a malicious, impermissible killing by Protagonist, and in both cases, the other five victims are maliciously killed.

If Five Killings Plus One is equivalent to Six Killings, and Six Killings is vastly dispreferable to Five Killings (as argued in the prior section), then Five Killings Plus One must be vastly dispreferable to Five Killings. And yet that results from just *one* additional murder.

The lesson is this: where the murder happens in the diagram matters. Bystander is concerned, in all of these cases, with the actions of Protagonist. If Protagonist murders someone, that matters a lot more than if some other murders take place that Protagonist might be able to prevent but that are not *her* murders. In other words, an extra murder by Protagonist has a completely different value for Bystander than an extra killing by some other agent whose killing might be affected by what Protagonist does (the comparison in line (4)). And this shows that there is no such thing as a "generic murder". *Who* murders matters.

In the final analysis, we see that it was a basic confusion for Chappell to criticize deontologists by appealing to the notion of the negative value to a Bystander of a "generic murder". There is no such thing in a sound deontology. Line (7) might seem to be true, but that illusion rests on the thought that Bystander counts killings without regard to who commits them. That, however, is precisely the assumption a robust deontologist rejects: robustness is the demand that Bystander care about what *Protagonist* does, not merely about how many killings occur.

Chappell defends his conception of preferability by describing it as (188):

> a theory-neutral conception of preferability. It begs no questions. It allows for features of actions (including their deontic status) to take priority over consequent value in determining which specific outcome (from among those that the agent can realize) one should overall prefer to see realized.

This is all true, as far as it goes. But it does not go far enough. It assumes that the two worlds can be compared "generically". That's not how a robust deontologist should frame value or preferability.

V. Conclusion

Chappell's argument fails for a rather mundane reason: he relies on strong transitivity, and it does not work in his deduction. But it also fails in two more interesting ways. First, it fails to recognize how strong deontological preferences for not performing acts that violate the rights of others are, even when doing so would prevent more such rights violations. Second, it presupposes that we can talk about "generic" murders. Deontologists should reject that idea for the same reason they reject wronging one to prevent others from being wronged: structure matters. We have to keep our eye on who has a claim on whom, and whose actions are being evaluated by bystanders.

Chappell concludes his article by claiming that (200):

> the following four features turn out to be mutually inconsistent:
>
> 1. Deontic constraints,
> 2. Robust normative authority,
> 3. Normative guidance, and

4. Adequate respect and concern for those who can be rescued at no further cost.[17]

He is mistaken. The patient-claim-grounded but agent-centered view we have described allows us to consistently represent all four features of a desirable morality. We accept deontic constraints, we give them robust normative authority, there are no choice points at which normative guidance is lacking, and we can still recognize that there are times when outcomes matter for what a Bystander should prefer and should, if possible, help bring about or prevent. What our view lacks is only the misguided idea that the value of worlds can be described with reference to "generic" murders.

With the idea of a "generic murder" Chappell snuck into his argument a basically consequentialist assumption about how a Bystander's preferences should work. We are not arguing here that consequentialism is wrong in its own terms. We are only arguing that he has failed to produce a new paradox for deontology. Deontologists can retain their commitment to respecting deontic constraints and view them as perfectly robust.

REFERENCES:

Alexander, Larry (2000). Deontology at the Threshold. *San Diego Law Review, 37*: 893–912.

Chappell, Richard Y. (2025). Preference and Prevention: A New Paradox of Deontology. *Free & Equal: A Journal of Ethics and Public Affairs, 1*: 175–201.

Chappell, Richard Y. (June 2025). The Curse of Deontology, *Good Thoughts*: n. 7, https://www.goodthoughts.blog/p/the-curse-of-deontology.

Nozick, Robert (1974). *Anarchy, State, and Utopia*. Basic Books.

Temkin, Larry (2012). *Rethinking the Good*. Oxford University Press.

[17] He clarifies the last point by adding in note 56 "That is, *after* one has already been, perhaps illicitly, killed as a means to initiate the rescue attempt."