

Shelly Kagans Kritik an Begründungen von Constraints

(aus: Shelly Kagan: *The Limits of Morality*, Oxford 1989, S. 24–32)

Let us, therefore, imagine a case in which it *would* promote the good to kill. Imagine that by killing one innocent person I can prevent the murders of two other innocent persons. That is, by violating the constraint against killing, I prevent two other violations of the same constraint, and this is the only way to promote this desirable outcome. In short, if I kill the one, I save the two. [...]

What the case does involve, of course, is a situation in which, by hypothesis, a greater amount of good will be done by killing than by refraining from killing. This description assumes, obviously, that it is worse – all other things being equal – if two innocent people die than if only one innocent person dies. But it would be hard to deny this claim. [...]

The moderate [s. c. deontologist] needs to defend a constraint against killing. Such a constraint would rule out killing even in cases where it would lead to the best consequences overall, so it should rule out killing the one to save the two. And this is just as the moderate would have it. He believes that it is morally forbidden for me to kill the one, even though this is the only way to save the two. But the moderate needs to defend this belief.

If I kill the one, I save the two. What reason can the moderate give for prohibiting me from doing this?

1. Moderates typically try to defend the constraint against killing by stressing what a horrible thing I am doing to the victim if I kill him. And it is indeed a horrible thing to be killed. But it does not seem that emphasizing this point brings the moderate any closer to justifying the constraint: The death of my victim is a horrible thing to happen to him; but the deaths of the two other potential victims are obviously horrible things to happen to *them*. My victim is uncompensated – but the two others would be uncompensated too. If my victim's death is a bad thing, then the deaths of the two would, together, be worse. Why should I be forbidden to minimize the badness?

As long as the moderate sticks to what happens to a person when the constraint is violated he gets no closer to justifying the constraint. Talk of the value of an individual's being able to shape his own life, the disvalue of a life being violated, and the like – none of this helps explain why a constraint against killing should be erected; it only supports the extremist's [s. c. the consequentialist's] view that it is permissible to inflict one such harm in order that a greater number of such harms – a greater overall amount of badness – may be avoided. The barrier to promoting the best outcome created by a constraint is inexplicable in terms of the badness of what happens to the victims.

2. The inadequacy of stressing the badness of what *happens* to the person when the constraint is violated may suggest shifting the focus of the moderate's argument to *the relationship which obtains between the agent and his victim*. Moderates have often emphasized the quality or character of the acts they wish to constrain – as distinct from merely noting their outcomes. If I kill my victim, I stand in a particular relationship to him – and moderates have stressed how horrible it is for that relationship to obtain. In killing my victim, even for a good end, I

am degrading him, violating his integrity as a person, treating him as a means, breaching our common humanity, and the like. Presumably I do not stand in these horrible relationships with all those whom I merely allow harm to befall.

Once again, however, the moderate's argument fails to bring him any closer to justifying the constraint. For however horrible the interpersonal relationship in which I stand to my victim, the relationships between the two would-be murderers and their victims would be equally horrible. Violating the constraint may put me in a horrible relationship with my victim, but if it prevents a greater number of violations then the moderate has as yet no reason to forbid it.

So far the moderate has concentrated on badness – the badness of what happens, and the badness of relations – but it seems clear that such an approach cannot justify a constraint against killing. For *badness* relates to the objective value of the outcomes. The extremist wants to promote the most valuable outcome open to him. Considerations of the disvalue of this or that can only affect judgments about what is the best outcome – but it can't possibly justify creating a barrier to promoting the best outcome once we determine what that is. Constraints cannot be erected on considerations of badness.

Perhaps when the moderate says how horrible it is to kill, we misunderstand him if we take him to be arguing for a constraint on the basis of the objective disvalue of killings. It may be so clear that considerations of disvalue are not sufficient justification for constraints, that it is unfair to read the moderate in this way. Perhaps his saying that it is bad to kill is simply a way of saying that it is *wrong* to kill – that killing is forbidden. But this, of course, is simply to state the claim, not to justify it; and surely, the moderate thought he was offering *reasons* for his view. So let us ask again: why is there a constraint against doing harm even when this is necessary to prevent greater harm?

3. We soon come to a fundamental intuition for the moderate. He feels there is something deeply wrong about the extremist's willingness to do evil in order to prevent it. Somehow it seems utterly against the spirit of morality to be flirting with evil in the name of the good. The extremist, of course, may bemoan the fact that he was forced to do evil at all; but the moderate's intuition is that we have *misconceived* our relation to evil, if we will enter into partnership with it for whatever purpose. It is the nature of evil that we should *avoid* it. It is the nature of evil that we should simply never turn our mind to bringing any about.

No doubt all of us have felt the force of this view. Indeed it is embedded in our language, making it all the harder to recognize it as an assumption which needs to be questioned. We may find ourselves being reassured by the thought that if it is *evil* to perform some act, then clearly one cannot perform that act. One cannot do an act which is itself evil – a fortiori one cannot do the evil act for some desirable end. Hence one cannot do evil – even that good may come of it. But such thoughts conceal a non sequitur; we cannot move without argument from 'one cannot do what is forbidden' to 'one cannot do harm'. The non sequitur is covered by the fact that both claims can be expressed as 'one cannot do evil' – and the ambiguity is doubtless due to the belief of ordinary morality that doing harm *is* forbidden, even for good ends. But the moderate must *show* that it is evil (forbidden) to do evil (harm). The extremist believes that although we cannot eliminate suffering from the world, we should at least do what we can

to minimize it. To the moderate, this attitude is fundamentally misguided. But whatever the force of his intuition, the moderate still needs to justify it.

4. The moderate may return to the nature of the relation between me and my victim. He may suggest that to deliberately harm another degrades not only the victim of the harm, but the agent of the harm. It is hard to escape metaphor here: when I kill another, I become *polluted*. (Even those who believe in thresholds for constraints will feel soiled and profaned if forced to do harm in order to prevent some catastrophe.) The constraint against killing – the moderate argues – respects the agent's need to preserve his moral integrity.

It seems, however, that a concern with the moral integrity of the agent cannot provide a justification for having a *constraint* against doing harm. At the very best it could only help ground an option to refrain from harming others in pursuit of the good; but if an agent were willing to sacrifice his moral integrity in order to promote the greater good, the moderate has given no reason to prohibit his doing so. If I am *willing* to kill the one out of my concern for the two, then the moderate still needs to provide a reason to forbid the killing.

The moderate might reply that even if this objection is sound and a concern with the moral integrity of the agent cannot justify a constraint – the extremist has been defeated nonetheless. Granted, some other argument will be needed to establish constraints; and granted, the moderate ultimately hopes to establish a far broader range of options than the mere option to refrain from doing harm; but still: if the appeal to moral integrity is indeed sufficient to ground even this limited option, then the extremist has been refuted, since he denies the existence of any options at all.

The extremist, however, has a compelling response. It is clear that no argument based on the loss of moral integrity can succeed – regardless of whether it is meant to establish a constraint or an option – unless killing the one does actually involve a sacrifice of the agent's integrity. But if killing the one to save the two is justified – as the extremist believes – then it is no sacrifice of moral integrity for the agent to do so. What this means is that the moderate's appeal to moral integrity simply begs the question against the extremist. The moderate believes that killing the one is morally forbidden, but he has still given no reason to accept this claim.

5. The moderate may return to the relation between agent and victim once again, and suggest that if I kill the one to save the two, I am failing to respect the one as a person. As a member of the moral community, each person deserves to be treated in accordance with his standing as a person; that is, I must treat other individuals as the persons that they are. But to deliberately harm an individual, even for a desirable end, is to reduce that individual to a mere means in my treatment of him, and to fail to accord to him the respect due him as a person. A person must be treated as an end, not a means – and so I am forbidden to kill the one.

The general tenor of this argument is, of course, a familiar one.

The concept of treating an individual as a person (or an end) is often used to try to bridge the gaps which have bothered the extremist. But the concept won't do the job. The extremist will certainly want to agree that each person deserves to be given the sort of treatment befitting the valuable type of being that he is. But the moderate is too hasty in assuming that the

appropriate sort of treatment rules out harming one person to prevent harm to others – for this is the very claim that he needs to establish. The moderate hasn't shown that treating a person with the respect that is due him requires something *different* from taking his well-being equally into account in weighing the objective value of an outcome and promoting the best outcome overall. To kill without justification is undoubtedly to fail to treat the person with respect, and the moderate *claims* that to kill one in order to save two is *not* adequate justification; but he has given no reason to think so.

The cataloguing of attempts by the moderate to establish the existence of a constraint against killing could go on; but I know of none that succeed. What is striking about all of these arguments is that they suffer not from failures of detail, but from failure to get off the ground altogether. The moderate's belief in constraints may well be backed by intuition, but I have argued that this is not sufficient to justify the belief if it cannot be supported by a coherent explanation. Not only has the moderate failed to give an adequate explanation; he has failed to give even the beginnings of one. (Shelly Kagan: *The Limits of Morality*, Oxford 1989, S. 24–32)