THE TRAP'S DILEMMA

David Lewis

The Bicentennial year is a fit time to recall an early contribution to 'regional philosophy'.[1] In the year 1879 Edward Kelly put forward this ingenious argument.

> I would like to know what business an honest man would have in the police. A man that knows nothing about roguery would never enter the force and take an oath to arrest brother, sister, father, or mother if required and to have a case and conviction if possible. Any man knows it is possible to swear a lie. And if a policeman loses a conviction for the sake of [not] swearing a lie he has broken his oath. Therefore he is a perjurer either way.[2]

At first glance, Kelly's example seems to fit right into present-day discussion of moral dilemmas, as follows. If the unfortunate policeman has taken an oath that obligates him to swear a lie under certain circumstances, and if those circumstances arise, then he has no right course of action. Either he takes a second oath to tell the whole truth and nothing but, and then he breaks it by lying; or else he doesn't, and thereby breaks his first oath to do everything possible to secure a conviction. Kelly's conclusion also looks familiar: it is because of his previous wrongdoing that the policeman afterward has no right course of action. An honest man would never have taken the first oath.

I think this first glance is misleading. Kelly's example is different, in two important ways, from the run of present-day examples. That is why a scrap of barroom wit deserves attention in this journal. However I shall leave the two differences in abeyance, and first ask the inevitable question: does Kelly's conclusion follow?

A man can be honest at one time and not at another. Kelly's conclusion seems to be that no policeman was honest when he entered the force. That conclusion does not follow. If an honest man is one who shuns oath-breaking, then it is safe to say that an honest man would not take the first oath if he thought it certain, or highly probable, that to fulfill his first oath he would have to take and break a second. But must an honest man foresee and shun even the slightest risk of being forced into oath-breaking? Suppose first that

---

[1] See Richard Sylvan, 'Prospects for Regional Philosophies in Australasia', *Australasian Journal of Philosophy* 63 (1985), pp. 188-204.

[2] Unpublished at the time; later published in Max Brown, *Australian Son* (Melbourne, 1948), p. 297, and elsewhere. Slightly copy-edited here.

the policeman foresaw a risk of dilemma, and estimated that the risk would be negligible — he was confident, wrongly as it turned out, that he would never need to swear a lie in order to secure a conviction. Or suppose he thought, contrary to what 'any man knows', that swearing a lie would not be possible; or anyway, that swearing a *convincing* lie would not be possible. Or suppose the policeman was none too clever when he took the first oath, and the idea that swearing a lie might someday be needed to secure a conviction just never crossed his mind. In the first case it seems that the policeman at the time of the first oath may have been, if not infinitely averse to oath-breaking, at least averse enough to count as honest. In the second and third cases, it seems there is no trace of dishonesty at the time of the first oath. An honest man can at least enter the force, if he is sufficiently optimistic or stupid.

But a harder, and interesting, question concerns honesty at the time of the *second* oath, if the policeman is eventually put to the test. Whether he was honest when he entered the force, years ago perhaps, is neither here nor there. Now that the only way not to break the first oath is to take and break a second oath, it is impossible not to be an oath-breaker. (And the policeman knows this.) Does that mean it is now impossible to be honest? We would think (1) that honesty is a trait of character, (2) that there can be no absolutely necessary connection between outer circumstances and inner character, and (3) that an honest man will never (knowingly) break an oath. But (1), (2), and (3) cannot be true together. We are forced to distinguish: to be *honest in character* is, in part, to be very averse to oath-breaking; to be *honest in deed* is, in part, never to (knowingly) break an oath. Given the distinction, it is hard to say which of the two best deserves the simple name: honesty. The surprise is that these are not the same thing, and furthermore that no amount of honesty in character is a certain guarantee of honesty in deed. Bad fortune is independent of honesty in character, perhaps, but it may necessitate dishonesty in deed.

The first big difference between Kelly's example and present-day discussion has now appeared. The present-day discussion mostly concerns dilemmas in which no course of action is right, or no alternative is good. What Kelly claims is different. He says that there is no way for the policeman to be honest. Kelly's example is a dilemma not in deontological or consequential ethics, but in the theory of the virtues.

Perhaps you've already thought that it wasn't much of a dilemma; because the policeman's first oath is morally null and void, either from the start or from the time it turned out to require him to swear a lie. (Maybe some part of its content retains some force, but not the part that makes the dilemma.) The easy answer to the non-problem about what course of action is right is that the policeman ought to break his first oath and testify truthfully or not at all. Yes indeed! But this easy answer about what's *right* does not even address Kelly's point about what's *honest*.

If you think rightness should be written into the very definition of honesty,

so that there's nothing dishonest about breaking an oath that would be wrong to fulfill, then you can ignore the difference between a moral dilemma about what's right and a dilemma about what's honest. It is an ancient idea that we can blur distinctions in this way. Myself, I think it a bad idea. In the first place it defies ordinary language: we can perfectly well say that sometimes it's wrong to do what's honest, so that even if there's nothing to do that's honest, still there may be something to do that's right. More important, if we blur the many virtues into some sort of nondescript overall rightness, we beg the question against a plurality of incommensurable values.

We may call Kelly's example a dilemma about 'obligation'; but only if we use the word in a strict and narrow sense, saving it for the sort of obligation that is undertaken by an oath, and not applying it indiscriminately to all manner of right conduct and pursuit of good. The policeman's misfortune is that his first oath, plus the circumstances, have put him under what we may call the *first obligation*: to do what is needed to secure a conviction. The first obligation requires him, for one thing, to take the second oath; and when he does, that will put him under what we may call the *second obligation*: to tell the whole truth and nothing but. The first obligation requires him, for another thing, to lie. So after he has taken the second oath, his dilemma is that he is under two opposite obligations arising from two different oaths. Such conflict of obligations is a familiar mainstay of present-day discussion.

But notice that even before he takes the second oath (if he ever does) the policeman already has a problem: whatever he does, he will break some oath, he will not be honest in deed, he will leave some obligation unfulfilled. And notice that his dilemma at this point is *not* a conflict of obligations. So far, his only obligation is the first obligation. He has no conflicting obligation. He is *going* to have a conflicting obligation, if he takes the second oath, but he doesn't have it yet. He is obligated to put himself under a conflicting obligation, but his obligation to become obligated is not yet an obligation *simpliciter*.

Here is the second interesting and unfamiliar feature of Kelly's example. If present-day discussion has led us to identify moral dilemmas with conflicts of obligations (insofar as they involve obligation at all) we are misled by the proverbial one-sided diet. It turns out that we have a different kind of dilemma: not a conflict between obligations in being, but rather an obligation in being versus an obligation to become obligated.

(You might say that even before he takes the second oath, the policeman already has an obligation to testify truthfully; this obligation conflicts with the first obligation; so we already have a conflict of obligations. I reply (1) that this looks like the nondescript sense of obligation that we wanted to set aside in order to concentrate not on what's right but what's honest; and (2) that we could stipulate that the prevailing customs are such that truthful testimony is not obligatory apart from the second oath; and (3) that even if there is a simple conflict between the first obligation and an obligation

to testify truthfully, this dilemma is not the same as the one that Kelly presents, even if the policeman is in both dilemmas at once.)

Kelly's example illustrates a little-known point about the interpretation of deontic logic.[3] A system of deontic logic has a sentential operator O which expresses deontic necessity. This is understood in the usual way in terms of accessible possible worlds: $O\phi$ is true at world w iff ϕ is true at all worlds deontically accessible from w. The accessible worlds are those which are in some sense 'ideal'; what exactly that means will vary from one intended interpretation to another. Also, we often impose a restriction to worlds in which certain fixed circumstances obtain: we have a certain time t in mind, and we limit ourselves to worlds that match the history of w up to t. For instance if t is the time when the policeman is choosing whether to take and then break the second oath, we thereby restrict ourselves to worlds where he has already taken the first oath, and where the circumstances are such that he can secure a conviction only by swearing a lie. Let us interpret O in such a way as to tie deontic necessity to the obligations (in the narrow sense) of the policeman: say that an ideal world is one at which, at time t and thereafter, the policeman never fails to fulfill any of his obligations. That means that, in view of the first obligation and the fixed circumstances, every accessible world is one where he takes and then breaks the second oath. But also, in view of the second obligation, no accessible world is one where he takes and then breaks the second oath. So there are no accessible worlds at all. That means that anything whatever is deontically necessary; $O\phi$ is true for any ϕ.

It is customary to read O as 'it is obligatory that . . .', especially when we have in mind the idealness that consists in perfect fulfillment of obligations. The surprise is that this customary reading for O is not quite right. It is deontically necessary to fulfill obligations that one is obligated to undertake but has not yet undertaken; but it is not yet obligatory to fulfill them. For instance in Kelly's example it is deontically necessary that the policeman tell the whole truth and nothing but; however until he swears the second oath — which he may never do — that is not yet, strictly speaking, obligatory. What is deontically necessary thus exceeds what is obligatory. The customary reading is safe enough if we use it with care and understand its limitations — just as it's safe enough to go on reading the hook as 'implies', so long as we know the difference — but still it is not strictly and literally correct. Deontic logic is not, strictly speaking, the logic of obligation.

Princeton University