

2000

On the Socratic Maxim

Joseph Raz
Columbia Law School, jr159@columbia.edu

Follow this and additional works at: https://scholarship.law.columbia.edu/faculty_scholarship



Part of the [Law and Philosophy Commons](#)

Recommended Citation

Joseph Raz, *On the Socratic Maxim*, 75 NOTRE DAME L. REV. 1797 (2000).
Available at: https://scholarship.law.columbia.edu/faculty_scholarship/753

This Article is brought to you for free and open access by the Faculty Publications at Scholarship Archive. It has been accepted for inclusion in Faculty Scholarship by an authorized administrator of Scholarship Archive. For more information, please contact scholarshiparchive@law.columbia.edu.

ON THE SOCRATIC MAXIM

*Joseph Raz**

Many years ago both John Finnis and I became interested in the Socratic view that it is better to suffer wrong than to do it.¹ My interest was triggered by Anselm Müller's lecture on the subject given at Balliol at that time.² Finnis discussed the issue in his *Fundamentals of Ethics*,³ where Müller's influence on him is acknowledged. At the time John Finnis and I debated the maxim and had a lengthy correspondence about it, but we did not convince each other. Now when I return to the issue, I can no longer remember the position I then took, except that I was the skeptic and Finnis the believer in the maxim. It is beyond doubt a dramatic and high-minded maxim. I remained intrigued by it and am glad of the opportunity to make this brief return visit. I should, however, declare at the outset that my interest is not historical. The different readings of the maxim and the different ideas associated with it in the discussion below are not brought forward as so many attempts to understand the historic Socrates or Plato. Rather, they are explorations of what role, if any, a maxim like this can have within a sound ethical outlook.

As before, I find the maxim's beguiling appeal in its ambiguities. It can be understood in a sense which makes it trivially true and in a way which makes it fairly clearly false. It can also be understood to be interestingly true, but not in a sense we can sensibly attribute to Socrates, nor one which expresses the high drama of Socrates's own choice when he refused to obey the Thirty Commissioners when they ordered him, on pain of death, to fetch Leon of Salamis to be executed. In what follows I will briefly explain a few different readings and the defects associated with them. I will explain why I remain a skeptic, why I feel that the rhetorical power of the maxim outstrips the insight it affords into the nature of morality.

* I am grateful to Penelope Bulloch and Ulrike Heuer for helpful comments on an earlier draft.

1 See PLATO, GORGIAS, 469c.

2 This lecture was published as Anselm Müller, *Radical Subjectivity: Morality versus Utilitarianism*, 19 *RATIO* 115 (1977).

3 JOHN FINNIS, *FUNDAMENTALS OF ETHICS* 112–20 (1983).

I. REFUTING EVENTISM

Finnis believes that the Socratic maxim is true. He also believes that it refutes a popular approach to ethics. Following Müller he calls that approach "eventism." Simply put, eventism is the view that the full moral relevance of an agent's action can be captured by the properties of the state of affairs that consists of his having so acted or by the event that the action constitutes.⁴ "Eventism" is an aspect of the approach commonly known as consequentialism. The refuting argument is meant to show that "eventism" cannot offer a reading of the Socratic maxim which does not make it self-contradictory. Since the maxim is, according to Finnis, true, it is not self-contradictory. Therefore, a view which entails that the maxim is self-contradictory is incorrect. Even those who doubt the Socratic maxim may accept that, if "eventism" cannot yield a non-contradictory reading of the maxim, that raises a serious doubt about it. The maxim does not appear to be contradictory or incoherent.

The following is one way of rendering Finnis's argument:⁵ for any two people *A* and *B*, let "the first event" (or "*First*" for short) design-

4 See *id.* at 113–14.

5 Finnis's original argument is as follows:

No eventistic ethics can accommodate (i.e., is compatible with) the Socratic principle. For any ethics that is eventistic will treat the following two formulations as equivalent:

- (1) It is true of every person *x* that *x* should act on the principle that "I ought not to do *V*."
- (2) It is true of every person *x* that *x* not doing *V* is a proportionately better state of affairs [event] than *x* doing *V*.

... The non-equivalence of (1) and (2) can be rigorously demonstrated in relation to at least one principle or rule of choice: the Socratic principle. Let us formulate it as Plato's Socrates himself did, i.e. not, primarily, as a comparison between states of affairs as better and worse, more and less harmful, but rather as a principle of choice: "If it were necessary either to do wrong or to suffer it, I should choose to suffer rather than to do it." (*Gorgias* 469c)

Thus it is an instance of (1):

- (1a) It is true of every person *x* that *x* should act on the principle that "I ought to suffer wrong rather than do wrong."

Now an eventistic ethic would have to translate that as:

- (2a) It is true of every person *x* that *x* suffering wrong is a proportionately better state of affairs than *x* doing wrong.

But (2a) contains a hidden self-contradiction. For (2a) yields, for instance,

- (2b) It is true of any persons *x*, *y*, *z* that *x* suffering wrong from *y* is a better state of affairs than *x* doing wrong to *z*.

And from (2b) there can be derived not only

nate the event of *A* wronging *B* and "the second event" designate the event of *B* being wronged by *A*. Obviously "the first event" and "the second event" designate the same event: that is, the first and the second event are one and the same event differently described. Let "the third event" designate the event of *B* wronging *A*, and "the fourth event" the event of *A* being wronged by *B*. Again, the third and fourth events are one and the same event. So we have the two identities:

- (1) *First* (*A* wrongs *B*) is the same as *Second* (*B* wronged by *A*).
- (2) *Third* (*B* wrongs *A*) is the same as *Fourth* (*A* wronged by *B*).

The maxim says⁶ that the second event is better than the third, and the fourth is better than the first; that is,

- (3) *Second* (*B* wronged by *A*) is better than *Third* (*B* wrongs *A*)—Finnis's eventist reading of the maxim, and
- (4) *Fourth* (*A* wronged by *B*) is better than *First* (*A* wrongs *B*)—*ditto*.

Given (1), we can substitute *First* for *Second* in (3) and get

- (5) *First* (*A* wrongs *B*) is better than *Third* (*B* wrongs *A*).

Similarly, we can also get (by substitution into (4)),

- (6) *Third* (*B* wrongs *A*) is better than *First* (*A* wrongs *B*).

But (5) entails the negation of (6), and (6) entails the negation of (5). Therefore, (5) and (6) contradict each other, and therefore whatever entails them is false.

What entailed them was the "eventist" understanding of the maxim which is encapsulated in propositions (3) and (4). This understanding of the maxim must therefore be rejected. If there is no other "eventist" understanding of the maxim, then, given the assumption

(2c) *a* suffering wrong from *b* is a better state of affairs than *a* doing wrong to *b*

but also its opposite:

(2d) *b* suffering wrong from *a* is a better state of affairs than *b* doing wrong to *a*.

In other words, the rendering of Socrates' principle attempted by the eventist (*viz.*, (2a)) entails that in a given situation one and the same state of affairs is both "proportionately better" and "proportionately worse," *i.e.* both more and less desirable (or valuable or choice-worthy). But such contradiction is as intolerable in ethics as in any other intelligent activity. So . . . all other forms of eventism can find no place for Socrates' principle.

Id. at 114–15 (emphasis added).

6 Note that this interpretation of the maxim is the one Finnis introduces in order to refute eventism. Neither he nor I accept it as correct.

that the maxim is true, or at least that it is coherent, "eventism" itself must be false.

I think that all of the above is sound. The problem is that there is another "eventist" reading of the maxim, and one which is clearly superior to the above, quite independently of the fact that it avoids reducing the maxim to a contradiction. According to it, (3) is not the correct way to understand the maxim. Rather it should be replaced by:

(3') *Second*, inasmuch as it involves *B* being wronged by *A*, is better than *Third*, inasmuch as it involves *B* wronging *A*.

(4) should be replaced with:

(4') *Fourth*, inasmuch as it involves *A* being wronged by *B*, is better than *First*, inasmuch as it involves *A* wronging *B*.

(3') and (4') do not entail (5) and (6). They do not allow one to form any "all things considered" judgment about the relative merit of any two events. They only allow one to judge that certain properties of some events count for or against them, or that they make them, insofar as these properties go, better or worse than some other events. More specifically, from (3') and (4') we *cannot* derive the absurd:

(7) Inasmuch as *First* involves *A* wronging *B* and inasmuch as *Third* involves *B* wronging *A*, *First* is better than *Third*.

All we can derive is the unilluminating:

(8) Inasmuch as *First* involves *A* wronging *B* and *Third* does not, *Third* is better than *First*, and inasmuch as *Third* involves *B* wronging *A* and *First* does not, *First* is better than *Third*.

In other words, judgments that certain aspects of an event are valuable do not warrant the conclusion that the event itself (that is, that all things considered the event) has such a value.⁷ Nothing in the Socratic maxim entitles us, on this interpretation of it, to form "all things considered" judgments as to the comparative merit of the two events. If that is an unhappy outcome, it is not the fault of "eventism." Which of the two events—*A* wronging *B* or *B* wronging *A*—is better is not something the maxim tells us about, or is meant to tell us about. That is true of "non-eventist" ways of reading it as well. The maxim does not tell us whether it is better for *A* to wrong *B* than for *B* to wrong *A*, only that it is better for *A* to suffer a wrong than to do it.⁸

7 This is, of course, the lesson, transferred here from reasons to values, taught by DONALD DAVIDSON, *How Is Weakness of the Will Possible?*, in *ESSAYS ON ACTIONS AND EVENTS* 21 (1980).

8 To remove any doubt, let me add that of course Finnis is fully aware of that last point.

II. AGENT-RELATIVITY

While the discussion above shows that the maxim is consistent with "eventism," it also points to another possible conclusion. The maxim may be inconsistent with agent-neutrality. It is not easy to give a non-trivializing characterization of agent-neutrality, and I will not try to be precise or exhaustive. Roughly speaking, agent-neutral views or theories are committed to the proposition that if two agents are in similar circumstances,⁹ then they face the same reasons, which can be described non-indexically. Agent-relative theories or views hold that some such reasons can only be indexically described. Thus, if a child is ill and two persons are in a situation where they can help him, then their reasons will differ if, for example, one of them is the parent of the child or otherwise related to him. Similarly, their reasons may depend on whether a result of their action will be a direct consequence of it or whether it will come about by the reaction of others to the action, or by the fact that in taking this action the agent omitted to take another, etcetera. Agent-neutrality implies that, other things being equal, it is right to kill in order to prevent others from killing a larger number of people, whereas agent-relativity is consistent with denying this assertion. My killing is my action and (perhaps) I should not undertake it "just" to minimize the number of killings in the world—killings which are not mine.¹⁰

With this rather vague explanation in mind, we can see that the maxim expresses a view consistent with agent-relativity, but not with agent-neutrality. How so? As explained in the previous Section, the maxim does not determine which event, *First* or *Fourth*, for example, is better. For a neutralist, for whom reasons for action depend on a perspective-free assessment of the value of relative possible states of the world, this would mean that the maxim does not establish how agents should act. But that is precisely what the maxim does. It determines that *A* should not wrong *B* and that *B* should not wrong *A*. They should not do so irrespective of the agent-neutral determination of the merit of certain actions. What matters to them is their own agency. *They* should not wrong others. They may or may not have reason to minimize the number of wrong-doings in the world. But if so, then that is separate from the reason they have not to wrong any-

9 "Circumstances" here include their own physical and psychological properties.

10 What if the other killings are mine? Suppose I can kill someone whose evil influence will make me kill several innocent people tomorrow. Does agent-relativity endorse such killing? Perhaps there can be several versions of agent-relativity, some relativizing to an agent and endorsing such killings, others relativizing to one agent at a time, and rejecting this.

one themselves. They know that they should not do that even if they do not know which act will minimize wrongdoing in the world.

I think that much of Finnis's discussion of the maxim is inspired by his feeling that it "refutes" agent-neutrality, rather than—as he claims—that it refutes "eventism." The compatibility of "eventism" with agent-relativity has been established by Amartya Sen in his *Rights and Agency*¹¹ in which he explained the structure of various agent-relative consequentialist positions. But while agent-neutrality may have been Finnis's real target, it is not obvious that the maxim does indeed refute it. One doubt, for example, may arise out of consideration of the stringency of the maxim. Is it not possible that agent-neutral positions can allow for the Socratic maxim so long as it is overridable? Perhaps the maxim means that the commission of a wrong by *anyone* is bad and should be avoided or minimized. If so, then one should not do wrong, but one should also minimize wrongdoing by others. Does that mean that it is better to suffer wrong than to do it? Of course it does. There is, after all, nothing wrong in suffering wrong, however unfortunate it may be. The Socratic maxim is true, but tautologically so. It is tautologically wrong to do wrong. A person who does wrong is acting wrongly, again tautologically. A person who suffers a wrong is not acting wrongly (that is, his suffering wrong does not entail that he does). Ergo, insofar as only facts mentioned in the maxim are concerned, it is clearly the case that it is better that one should suffer the wrong than do it.

This is (a version of) the trivializing reading of the maxim that Finnis criticized. There is no doubt that it is not true to the spirit of the maxim, and we shall return to the reasons for this below. Interestingly, however, even the trivialized reading of the maxim is inconsistent with agent-neutrality. How so?

Some wrongdoings may well be worse than others. If forced to act wrongly (assuming, as I do, that such a situation is possible), one should choose the lesser wrongdoing. A neutralist will also say that one has reason to reduce wrongdoing in the world, and therefore one may commit a lesser wrong to stop others from committing a worse one (the same assumption needs to be invoked to allow for that possibility). The Socratic maxim is silent on the relative wrongness of different wrongdoings. This may appear natural, because the maxim compares not two wrongdoings but a doing and a suffering of a wrong. However, there is no wrong suffered unless there is a wrong done. If I should suffer a wrong rather than do one, it follows that I should not do a wrong to stop another from wrongdoing (in this case

11 Amartya Sen, *Rights and Agency*, 11 PHIL. & PUB. AFF. 3 (1982).

from wronging me). Is this always sensible? Suppose his wrongdoing is worse than the one I would commit in stopping him—should I wrong him slightly to stop him from greatly wronging me? The maxim is best understood as being silent on this question.¹² Let us assume that the maxim applies only where the wrong suffered and the wrong needed to avert it are roughly the same. (Alternatively, let us focus just on the cases where this is so.) A neutralist is committed to the view that, other things being equal, it does not matter whether I commit a wrong to stop another from wronging, or whether I do not commit any wrong and the other person does. For the neutralists all that matters is that the number of wrongdoings will be as small as possible. They do not care who are the wrongdoers, myself or others. The maxim is clearly not indifferent in this sense. Rather, it holds that I should not commit a wrong—that if I do not then someone else will is neither here nor there.

The maxim is agent-relative. But does it refute the agent-neutral approach to ethics? This is far from clear. This argument is unlike Finnis's own argument, which aims to derive a contradiction. Were Finnis's argument successful, it would have established the falsity of "eventism" on the assumption that the maxim is coherent. This time the most the argument can hope to establish is the incompatibility of the maxim and agent-neutrality. Why should that be an argument against agent-neutrality? Things are even worse for anyone who seeks to use the maxim to refute agent-neutrality. While the maxim itself is agent-relative, it is derivable (perhaps in a slightly weakened form) from agent-neutral premises. After all, if all you care about is to reduce wrongdoings by whomever, and your choice is between doing wrong to stop another from wrongdoing or refraining from doing wrong yourself, you should—barring exceptional circumstances—choose the second option, and not do wrong yourself. You will then know for certain that that wrongdoing will not occur. There is always a risk that by doing wrong to stop another from wrongdoing you will fail to achieve your goal, or that your wrongdoing will be unnecessary because the other may fail to do the wrong he intends or may even change his mind about the action (not necessarily to avoid wrongdoing).

It is doubtful that the maxim can be used to refute "eventism"—Finnis's declared aim—or to cast doubt about agent-neutrality—which may have been his real aim.

¹² I will return below to the question of the coherence of the situations described herein.

III. SELF-INTEREST AND BEING WRONGED

It is time to return to the fact that we still do not have a credible interpretation of the maxim. This too is one of Finnis's points. He discusses the ways of understanding the maxim which were relied on above but rejects them as inadequate.¹³ My reason for rejecting them is, however, somewhat different. The trivializing reading, for example, misses the point of the maxim. It simply reasserts that morally speaking what counts is not doing wrong. The maxim, however, bridges the gap between morality and personal interest. It says that it is better for the agent to suffer wrong than to do it. That is, it does not assert that suffering wrong is morally better, but rather that, from the point of view of the agent's own interest, it is better to suffer wrong than to do it.

Herein lies the dramatic force of the maxim. It is common to think that wrongdoing is bad for the wronged person, but not for the wrongdoer. Socrates (just about) reverses it all: wrongdoing is worse for the wrongdoer than suffering wrong. Is he right? A doubt is raised by the very sweep of the assertion. After all, some wrongs are much worse than others. Could it be that it is worse for the agent to do the slightest wrong than to suffer the greatest one?

You may think that it is impossible for such a choice to arise. If a big wrong can be averted by committing what would otherwise be a minor wrong, then it would not be wrong to perform the act which in other circumstances would be a minor wrong. Something like the legal defense of necessity will apply and turn it into a permissible act. Therefore, if the choice is between doing a wrong, which will avert a major wrong, and suffering that wrong, then (if it is to be wrong to do it) that wrong must be at least as great as the wrong it will avert.

But this reply invites two responses. First, the maxim as stated is not confined to doing wrong to avert the wrong one would otherwise suffer. It compares doing and suffering wrong in complete generality. Second, even when the wrongdoing is undertaken to avert the wrong one would otherwise suffer, it does not follow that it is at least as great. In self-defense we are, perhaps, justified in stopping or trying to stop an aggressor by actions which otherwise would be as great a wrong as the wrong he attempts against us, or even greater. This, at any rate, is the common opinion. But self-defense is directed against the would-be wrongdoer. Necessity can apply to acts that affect innocent parties. They are not responsible for the wrong we are trying to avert, and acts which wrong them normally remain wrongful acts even when needed to avert a greater wrong against oneself. Sometimes such wrongdoing

13 See FINNIS, *supra* note 3, at 112-120.

is permissible, but it remains wrongdoing and requires its perpetrator to apologize to the innocent third parties and compensate them.

Perhaps these replies can be averted. Perhaps the maxim should be confined to comparing doing wrong to avert wrong to oneself and/or not taking the wrongful action and suffering that wrong. Perhaps whenever necessity makes an act permissible, it is not wrongful action. Puzzles arising out of the sweep of the maxim remain. Its dramatic flourish suggests that to be wronged is against one's interest, that it is bad for one. But often it is not. Often, when people break duties they owe me, while they wrong me, they do me no harm. Someone may promise to wake me up early in the morning so that I will not miss a plane. He fails to do so. This is wrong of him, and he wrongs me in doing so. But as it happens, I woke up in good time and comfortably caught the plane. Someone libels me, spreading malicious falsehoods about me. He wrongs me, but may cause me no harm since, say, no one believes him. Someone denies me a job offer to which I am entitled, and which he has a duty to me to offer me (there may be, for example, some fiduciary relationship between us, obliging him to offer me the job). He wrongs me, but I may not suffer as I may have a better job offer that I would have preferred anyway. Examples of such cases are plentiful. Moreover, sometimes being wronged turns out, sometimes predictably, to be in one's interest. It may satisfy a psychological need to be a victim. It may make one the center of welcome attention and sympathy, and so on. In all such cases it sounds falsely melodramatic to claim "I would rather suffer such wrongs than commit a wrong." Of course, what else would you expect?

Perhaps this point mistakes Socrates's meaning or rather (since I do not think that that was Socrates's own meaning) mistakes the meaning it should be understood to have. The maxim does not mean that being wronged causes one harm. It means that being wronged is itself a harm which one suffers, is itself against one's interest. But nevertheless it is better to suffer it than to commit a wrong. But that notion is obscure to me, and I will not try to explore it here. If suffering wrong is not in and of itself bad to the wronged person, the force of the maxim is undermined. Of course I would rather be wronged. No harm in that. If the wrong brings harm in its wake, maybe I would rather do wrong than suffer that harm, but that is because of the harm which will befall me, not because it is the result of a wrong done to me. I would take the same wrong action to avoid a harm caused by an earthquake or other act of God or of chance.

Finally, there is the problem of self-sacrifice. If it is always better to be wronged than to do wrong, and that from the point of view of

the agent's own self-interest, it would seem that it is impossible for people to act against their own interest in order to do what is morally required of them. Whenever one is morally required to act in a certain way, not doing so is morally wrong, and therefore, by the maxim, against one's interest. Therefore, one is always acting in one's own interest when doing what one is required to do. This argument assumes that following morality will not be even more against one's interest than flouting morality. This assumption is not entailed by the maxim but seems to follow from its spirit. But acting against one's interest in a moral cause is possible and sometimes required.¹⁴ If the maxim has the consequence of denying this, it must be rejected.

IV. WRONGDOING AND THE AGENT'S INTEREST

I may be too literal-minded. I may take the maxim at its word when I should realize that it is stated as it is for dramatic effect, whereas its point, real enough, has nothing to do with the rhetorical comparison between wrongdoing and suffering wrong. Its point is simply to assert that wrongdoing is bad for the wrongdoer.

If all that that means is that *sometimes* the very act of wrongdoing is bad for the agent, I have no quarrel with the maxim. But if it means that the very act of wrongdoing is *always* bad for the agent, doubts arise. They arise simply because some acts of wrongdoing are fairly trivial, and wrong though they are, they may not affect the agent at all. Imagine an inadvertent disclosure of confidential information which is in fact of no consequence. One acts wrongly in betraying another's confidence, and the action is culpable, for one should have been careful to avoid the disclosure. But the matter is of no consequence to the person who is so betrayed, nor to anyone else.¹⁵ Is it bad for the agent? I assume that if it is bad, that is because it has some bad consequences for him, for example, undermining other people's trust in him, or because it either detracts from his well-being or mars his character. But if it does either, it cannot be a totally trivial and forgettable action. Some wrongdoings are, however, totally trivial and forgettable. Therefore, it is not true that all wrongdoings are bad for their perpetrators, even though they often are. Perhaps the dramatic note struck by the Socratic maxim includes a good deal of rhetorical exaggeration.

14 See my discussion of the possibility of such sacrifices in JOSEPH RAZ, *ENGAGING REASON* 315ff (2000).

15 Is betrayal, however trivial and inadvertent, bad in itself to the betrayed? I tend to feel that it is bad in itself only to the extent that it expresses an inappropriate attitude by the betrayer. This, however, is not always the case. Betrayal, as most other actions—including wrongful ones—can be due to bad luck.