

Socrates Defending Himself: Never Return Harm for Harm-- Crito 49cf

Dr. James Butler, Ph.D.

© 2001

Site Managers Note: This paper was written in partial satisfaction for the rank of shodan in Waboku Jujitsu. The opinions expressed in this paper are solely those of the author and do not necessarily represent the opinions of the Waboku Jujitsu Group, its members, or lackeys.

Arguably one of the greatest philosophers of all time is Socrates of Athens (469-399 BCE). He not only dedicated his life to philosophical inquiry, but he inspired many others (Plato and Aristotle being the two most prominent) toward a more ethical life. Yet, Socrates is as much an enigma as he is a hero; all we know of him comes to us second-hand. Socrates wrote nothing, hence we must piece together the features of his life and thought from, among other testimony, the early dialogues of Plato, his most famous disciple.

Broadly speaking, this paper addresses one aspect of Socratic thought that we find in Plato's dialogues, namely Socrates' rejection of retaliation. Socrates claims that one ought never to return injustice for injustice or harm for harm. Part of his reasoning for this is that he claims, somewhat surprisingly, that 'there is no difference between doing someone an injustice and harming him/her' (Crito 49c). There seems to be, however, a clear counter-example, namely physically defending oneself from attack. For it seems that self-defense harms the attacker, but it is not unjust. The motive for saving Socrates' position from this alleged counter-example is not simply because it is an intellectual puzzle; it has bearing on how we understand Socrates the person and how this great mind understood the concept of self-defense. Surely Socrates was forced to deal with such situations; it is well-documented that Socrates served as a footsoldier in the Athenian Army. If we assume that Socrates, like any other soldier used a weapon against Athens' enemies on the battlefield, those military actions appear to be in conflict with his long-held belief that one should 'never harm another person' (Crito 49e). How is it, then, that Socrates the soldier does no harm to members of the oncoming army?

By examining passages in the early, 'Socratic' dialogues of Plato where Socrates expounds his notion of harm, I will show that, in opposition to most modern notions of harm, Socrates believes that in a case of justified self-defense (such as a military battle), one would be doing no harm to an attacker by physically wounding him or even killing him. Just as a doctor doing triage must amputate a limb for the greater good of the patient (even without the patient's consent), Socrates believes that physically preventing the attack from succeeding provides the *attacker* with a greater good, namely the preservation of the attacker's psyche.

§1

Socrates argues in the Crito that even though he was wrongly convicted, he ought not escape from prison. For if he did escape, he would be unjustly breaking the covenant he has with the laws of Athens. A key element in his argument is the principle that:

(a) one ought never to return injustice for injustice. (49b10-11)

Socrates, however, makes a surprising addition: He states that because *there is no difference between doing someone an injustice and harming him/her* (49c), it follows that

(b) one ought never to return harm for harm. (49c4-5).

Let's call (a) and (b) 'Socrates' two principles of non-retaliation'. Socrates evidently thinks the two principles are identical, since he says that there is no difference between doing an injustice and doing harm.¹

¹Following Penner ("Is the Crito a Treatment of Political Obligation?" unpublished manuscript p.15), I take Socrates statement that there is no difference between the two principles as an identity statement.

But why should we think, like Socrates, that there is no difference between doing harm and doing injustice and thus the two principles of non-retaliation amount to the same thing? In the case of self-defense at least, the two principles of non-retaliation seem to give markedly different answers. Suppose you are innocently standing on the sidewalk and a man attacks you with a knife, attempting to stab you. Let us suppose that, because you are innocent, the perpetrator is trying to do an injustice to you as well as harm you.² If you choose to use physical force to defend your life -- say by breaking the perpetrator's arm thereby making him drop the knife -- it does not seem that you are doing anything unjust. For, we normally think that there is nothing unjust about defending one's life, even at the expense of the attacker's life. Thus, if you are not committing an injustice by physically defending your life, this self-defense scenario is simply not a case of *returning injustice* for injustice.

By contrast, the other principle of non-retaliation -- never returning harm for harm -- seems quite applicable to the self-defense case. If, in order to avoid being harmed (stabbed), I break the perpetrators' arm, thereby making him drop the knife, it certainly seems that I have harmed the perpetrator. After all, breaking someone's arm seems to be doing him/her physical injury.

Thus, the self-defense scenario seems to show that, contrary to Socrates' claim, there is a difference between doing harm and doing injustice; physically defending oneself seems just, yet harmful to the attacker. Is Socrates simply confused?

I propose to look at two interpretations of the Socratic principles of retaliation, both of which attempt to preserve the identity between doing injustice and doing harm in light of the self-defense example. The first concedes that using physical force on an attacker is harming him and since there is no difference between harming and doing injustice, defending oneself is, contrary to common opinion, unjust. This first interpretation I shall call the 'Jainist interpretation', after the Jainist principle of non-violence (ahimsa) which Gandhi is thought to have advocated.³ The second, which I will endorse, surprisingly argues that using physical force against an unjust attacker is, by the Socratic notion of harm, not harming him at all. In fact, if it is possible to harm another by *failing to do* some action, then in certain cases, the principle never return harm for harm might *necessitate* that one physically defend oneself. For example, if an attacker is attempting to do an injustice to me, not only am I *not* harming him by physically defending myself, I may in fact be harming the attacker if I *do not* defend myself. Let us call this second interpretation the 'Just Force' interpretation.

§2

Background

Before looking at the two interpretations which attempt to preserve the identity between harm and injustice, we need to investigate the Socratic notion of harm. For until we understand what the Socratic notion of harm is, we shall not be able to understand why he identifies doing harm with doing injustice. The investigation will primarily rely on passages in the Crito and Apology where Socrates makes two distinct comparisons: (i) the injury suffered from killing unjustly versus the injury suffered from being killed unjustly, and (ii) the importance of the body versus the psyche in living a good life. We shall also look at a passage in the Protagoras where Socrates claims that humans are naturally determined to avoid harm.

Unfortunately, Socrates' account of harm is not easily discerned. As is all too common with Socrates, he never presents a comprehensive theory; we are only presented with clues in different dialogues. We get our first glimpse of what harm is when Socrates argues with Meletus about corrupting the young in the Apology. Socrates' argument runs as follows:

1. Bad people do harmful things to their closest neighbors(25c5-9)
2. No one wants to be injured (25d3)

²We are assuming for the sake of argument that the attacker is doing an injustice. Clearly, in the real world such assumptions are not so forthright. Oftentimes, we may think someone is doing an injustice, but we are simply wrong. But for the sake of the argument, let us assume there are no epistemological problems in truly believing that the attack on one's person is unjust and harmful.

³The term Jainist comes from the Indian philosophy, Jainism, of whom Gandhi was somewhat of a disciple. They claim that one should *never* use violence against another living creature.

3. If Socrates corrupts those around him, he will likely be harmed in return by those corrupted. (25e1-4).⁴

Thus,

4. Socrates either does not corrupt, or because corrupting leads to getting himself harmed (which he does not want), he corrupts unwillingly. (25e4-5)

The important premise here for our purposes is that Socrates claims that no one wants to be injured (or harmed).⁵ If true, what does this tell us about what harm is, according to Socrates? Surely, harm cannot be equivalent to any short-term pain or suffering; for it is quite common for people to want to endure short-term pain or suffering in order to gain some long-term benefit. For instance, I want the painful tetanus shot so that I will not become sick from stepping on a nail. Do we say that the shot 'harms' me because it is painful? Certainly not. The long term benefits of not becoming sick are such that, though painful, the shot is actually a benefit.⁶

So, if people do not want to be harmed, and harm cannot be short-term pain or suffering, what might Socrates' idea of harm be? It seems, from the example of the tetanus shot, that Socrates' notion of harm must take into account the long-term consequences, as well as some notion of 'net' harm or 'net' happiness. Yet, how far does this desire not to be harmed extend? In the Protagoras, Socrates suggests that one wants the maximum amount of benefit and the minimum amount of harm:

No one who either knows or believes that there is another course of action better than the one he is following will ever continue on his present course when he might choose the better. To 'act beneath yourself' is pure ignorance; to 'be your own master' is wisdom...Then it must follow that no one willingly goes to meet harm or what he believes to be harm. To make for what one believes to be a harm, instead of making for the good, is not, it seems, in human nature, and when faced with the choice of two harms no one will choose the greater when he might choose the less. (Prot. 358b-d).⁷

It is a bit unclear what this passage tells us about what counts as a harm. For the scope of 'harm' depends upon Socrates' inference ('it must follow') from one must choose the (apparently) best available course to 'no one willingly goes to meet harm or what he believes to be harm'. There are, however, two possible interpretations of 'harm' which would satisfy Socrates' inference:

The first inference, relying on the fact that best and worst are contraries, implies that since we must choose the (apparently) best option, it is impossible to choose the worst option, though there may be other, medial options available. Let us call this the 'limited inference'. The second inference, on the other hand, suggests that since one must choose the (apparently) best option, one could not choose *any* option which is inferior to the best available. Let us call this the 'expanded inference'.

Whether Socrates chooses the limited or expanded inference will have critical implications on what he takes a 'harm' to be. By the limited inference, only the worst available option need be a harm; other medial options cannot be chosen but need not be thought of as harms. This limited inference reflects the common view that there are benefits, harms, and those things which are neither harmful nor beneficial. The expanded inference, on the other hand, makes any option less than optimum a harm. Since

⁴Socrates seems to take 'harm' as equivalent to 'injure', for the argument turns on the idea that if one is harmed, one does not get what one wants (i.e. not to be injured).

⁵Although there are other very interesting philosophical issues embedded in this argument-- for example, the theory that I do not desire something that LEADS to me getting harmed (see Gr.468c)-- I cannot address them all here. For this paper, I limit the discussion to what Socrates' notion of harm is.

⁶A similar argument can be constructed to show that benefit cannot be short-term pleasure; many of the actions we perform which are beneficial to us (surgery, cavities filled) are not pleasant in the short-term. To see how such an argument might run see Gorgias 468.

⁷ Translated by W.K.C. Guthrie in Collected Dialogues p.348-9.

we cannot make for and medial options -- even though they may not be the worst options available -- and since Socrates claims that we cannot make for a harm, these medial options would be harms.

Which type of inference does Socrates draw? He endorses, I believe, the 'expanded inference': any option short of optimal is a harm, even in the case where all options seem to benefit the agent. This can be seen if we look at the opposite case, where every option is detrimental: According to his statement in the *Protagoras*, one must both (i) choose the (apparent) good and (ii) choose the (apparent) lesser of 'two harms'.⁸ Even though the choice between 'two harms' (say chemotherapy or cancer) is far from ideal, it is *human nature* to make for the (apparently) *good* option. Notice that in this scenario Socrates never distinguishes between *what is the lesser harm* and *the good*: It is reasonable to suppose, then, that the least injurious option *simply is the good* in such unfortunate circumstances,.

We should, then, expect parity of reasoning in the opposite case: the choice of two 'goods' (i.e. two options which are somewhat beneficial). Since in the case of 'two harms' Socrates calls the least injurious option 'the good', in the case of 'two goods', the lesser of the two would be a harm, even though it might be somewhat beneficial to me.

Even with this reasoning, it might seem counter-intuitive to suggest that something which is beneficial (though not maximally) to my happiness is a harm. Take the following example however: In order to benefit myself over the long run, I should have the cavities in my mouth filled. Nowadays, I have two options for filling my teeth: I can (i) have a painless laser filling or (ii) have the old-fashioned drilling. Both options would fix my teeth properly and contribute positively to my happiness for the rest of my life. Yet, if I were to choose the second best option available (the old-fashioned drill and fill), I would suffer the pain of the drill needlessly. Such a choice, assuming that I knew of the laser filling, would be, as Socrates claims, pure ignorance. Thus, where both might appear to be 'goods', one turns out to be a harm because it causes me needless suffering. If this parity of reasoning is sound, then what constitutes a harm, according to Socrates, is not whether I am benefited in some way or other; harm turns out to be *any* course of action which is not the best in the present circumstances.

Combining this notion of harm with the *Apology's* suggestion that harm is concerned with long-term consequences upon one's happiness, we end up with the following comprehensive account of harm: a harm is the action which, in the present circumstances, fails to maximize one's happiness over a complete life.⁹ By such an account of harm, a tetanus shot would not be a harm because, given the alternative of avoiding the shot but getting sick, the shot maximizes one's future overall happiness. Failing to get the tetanus shot, on the other hand, would be a harm because by getting sick, one's future happiness is not maximized.

Notice that embedded in Socrates' account of harm is the notion that what makes an action harmful in a given situation depends on its consequences for one's future happiness. Thus, what might be harmful in one set of circumstances might not be a harm in a different set of circumstances. For instance, taking aspirin against some colds is advantageous to my health, but doing so against influenza can be detrimental.

⁸The situation in which one must choose between "two harms" needs to be interpreted in light of Socrates' statement that "to make for what one believes to be a harm instead of making for the good is not in human nature". If it were the case that both options are "harms", and one option must be chosen, Socrates would be wrong that humans cannot make for what is thought to be a harm. Indeed Socrates himself would be refuting his own claim about human nature! We should, therefore, take Socrates' claim about a "choice of two harms" figuratively. Namely, Socrates is pointing out that, in a situation where neither option is particularly appealing (e.g. chemotherapy or cancer), people still want the option that will maximize their happiness. And the option which does maximize one's happiness (out of the possible options) is a good, not a harm.

⁹Brickhouse and Smith (*Plato's Socrates* p.121-3) call this notion of harm "relative harm" which is contrasted with (what they believe to be) another Socratic notion of harm, "absolute harm". An absolute harm, on their view, is the only thing that is independently bad (i.e. bad in itself) which turns out to be vice. I disagree that Socrates needs (or has) such an absolute notion of "harm". For he can argue, using only the relative notion of harm, that vice is always a harm. That is, when compared to any other option, vice never maximizes one's happiness over a complete life.

Accordingly, when we speak of something as being an 'unconditional harm', we are, most often, speaking improperly; for being 'unconditional' suggests that in any circumstances, and compared to any alternatives, such actions are always worse than any other possibility.¹⁰ But depending on where I am now and the possibilities open to me, it is frequently the case that these so-called 'unconditional harms' are, or ought to be, chosen because they are beneficial. For example, it might be said that losing a leg is one of these 'unconditional harms' since no one really *wants* to lose a leg. But, to see if losing the leg is a harm, one would need to know the circumstances in which the leg is lost and compare it to all the other possible options. If the leg were lost in a car accident, losing the leg is most likely a harm; for its loss adversely affects one's future happiness compared to having the leg. If, however, one lost the leg in order to prevent the further spread of gangrene, then its loss is, like many other surgical techniques, a benefit, not a harm. So, on this view of harm, something painful or even devastating cannot be called a harm unless, given all the other possible options, it fails to maximize my happiness over a complete life from the present circumstances.¹¹

In essence, the Socratic view of harmful action, as I have presented it, suggests that, given a particular set of circumstances, one never wants (and thus prudentially never ought) to trade something more important with regard to their long-term happiness for something less important. Just as in the example of the gangrene leg above, one does not want, in these circumstances, to trade something more important to one's happiness (life) for something less important (a leg). And so in all cases: since one wants the best option, one must compare the various outcomes on long-term happiness and never trade something more important with regard to happiness for something less important. To make such a trade, would be ignorant and would harm oneself.

If Socrates' notion of harm requires one to compare various alternatives and their effect on happiness, let's look at some instances where Socrates compares the harm of the following: (i) killing unjustly versus being killed unjustly and (ii) the corruption of the body versus the corruption of the psyche. Regarding the first, in the Apology, Socrates claims:

...if you kill one such as I am, *you will injure yourselves more than you will injure me*. Nothing will injure me, not Meletus nor Anytus-- they cannot, for a bad man is not permitted to injure one better than himself. I do not deny that Anytus may, perhaps, kill him, or drive him into exile, or deprive him of his civil rights; and he may imagine, and others may imagine that he is inflicting a great harm upon him: but there I do not agree. For the harm of doing as he is doing-- of trying to unjustly take a man's life-- is far greater.(30c6-d5) [my italics].¹²

According to Socrates, one who unjustly kills is harmed more than one who is unjustly killed.¹³ But why is this true? Socrates position, I believe stems from his view of the relative worth of the psyche compared to the body. At Crito 47bff, Socrates asks Crito the following:

There is a part of us which is improved by healthy actions and ruined by unhealthy ones. If we spoil it by taking the advice of non-experts, will life be worth living once this part is ruined? The part I say is the body. Do you accept this?

¹⁰I say above that "unconditional harm" is *most likely* a bit of improper speech because there is in all likelihood one thing, vice, which will always fail to maximize one's happiness. See previous note.

¹¹It might be objected that since, according to Socratic doctrine, a harm is something that fails to maximize one's happiness over a *complete life*, one cannot judge until one is dead whether one choose something good or harmful. Actions like eating oat bran might be thought to be medically beneficial now, but it may be discovered in ten years that oat bran is found to cause cancer. Thus, we cannot assess the benefit or harm of eating oat bran until the consumer is already dead. All of this is not an objection against the truth of Socrates' *ethical theory*, however. It merely points out an epistemological problem we humans face regarding such an ethical theory.

¹²From Benjamin Jowett's translation in The Dialogues of Plato, Erich Segal ed., Bantam Books, 1986.

¹³Socrates defends this same view a Gorgias 469bff.

Yes.

Well is life worth living with a body which is worn out and ruined in health?

Certainly not.

What about that part of us which is mutilated by injustice and benefited by justice? Is life worth living with this part ruined? Or do we believe that this part of us, whatever it may be, in which justice and injustice operate is of less importance than the body?

Certainly not

Is it more precious?

Much more.

Although Socrates does not explicitly mention what is ruined by injustice, it seems reasonable to conclude, based on the contrast with the body, that he is referring to the psyche.¹⁴ Thus, if the psyche is ruined by injustice, and the psyche is more precious to us than the body, then the ruination of the psyche through injustice is more of a harm than the ruination of our body.

We now see why Socrates claims that Meletus and Anytus are harming themselves more than they are harming him by pursuing the death penalty. Socrates is only facing the ruination of his body, while keeping his psyche free from injustice. Meletus and Anytus, on the other hand, are ruining something more precious than their bodies, their psyches, by perpetrating an unjust act.

The Socratic account of harm, as I have interpreted it, then, might be summed up as follows: For Socrates, a harm, whether it be harming oneself or another, consists of failing to maximize that person's happiness over a complete life, given the present circumstances. Thus, if we are to determine whether or not something is a harm, we must compare all the possible options on our bodies and psyches in order to determine which actions fail to maximize happiness.

§3

Two Interpretations of the Self-Defense Case

From the above discussion of harm and the relative importance of the psyche in comparison to the body, I believe that there are two ways of interpreting the self-defense case which would preserve Socrates' claim that there is no difference between doing an injustice to someone and harming him/her. Both interpretations assume that IF one harms the attacker while physically defending oneself, then it is an injustice. The two interpretations differ, however, in their suppositions about whether or not one harms the attacker by physically defending oneself. The first, which I have called the 'Jainist interpretation', concedes that physically defending oneself is harming the attacker. The second interpretation (which I have called the 'Just Force' interpretation) denies that one who uses physical force against an unjust attacker is a harm at all.

The Jainist Interpretation of Non-retaliation

The Jainist interpretation defends the view that there is no difference between doing an injustice to someone and doing them harm by interpreting Socrates' two principles-- never return harm for harm and never return injustice for injustice-- along the lines of the Jainist theory that one ought never do any violence at all:

We...say, speak, assert and preach: 'All animals, living beings organisms, and sentient creatures should not be injured, governed, enslaved, tortured and killed: Know that it is non-violence which is (completely) free from sin'.....

Having discerned this, a sage should neither use any weapon causing violence to the mobile being, nor cause others to use it nor approve of others using it.¹⁵

¹⁴For a noticeable similarity to the Crito, see Protagoras 313aff where Socrates claims that the soul is much more valuable than the body.

¹⁵ From Ayaro, translated by Muni Mahendra Kumar, in Understanding Non-Western Philosophy, Daniel Bonevac and Stephen Phillips eds. Mayfield Publishing, 1993, p.141-142. It is worth noting that the last part of the Jainist passage directly condemns Socrates' alleged behavior on the battlefield. Socrates, I assume, used weapons on the battlefield, and used them against the enemies of Athens.

By condemning the use of all violence --which we will assume includes physical force against another-- the Jainist theory suggests that any physical force used against another creature is a harm and, therefore, is unjust.¹⁶ Thus, using physical force to defend oneself in the self-defense case would be harming the attacker, and contrary to appearances, would be an injustice. Socrates' identification of doing harm and doing injustice, therefore, is preserved in light of the self-defense case.

If we put this Jainist reading of Socratic principles into a more formal argument, it might run as follows:

1. Since the psyche is more important than the body, ruining the psyche (by doing injustice) is more harmful than ruining the body (Crito.47e).
2. Defending myself harms the attacker.
3. There is no difference between doing an injustice and doing a harm (Crito 49c).
By 1, 2, & 3
4. Doing an injustice to the attacker ruins my psyche (Crito.47e-48a) while not defending myself ruins my body (I get stabbed).
5. Since no one wants to be harmed (Ap. 25d), I ought¹⁷ not harm the attacker, preserve my psyche, and if necessary, sacrifice my body.

What textual evidence is there to support the Jainist interpretation? It might be argued that since Socrates, in his argument with Meletus (25c-e), makes no differentiation between 'harm' (*kakØn*) and 'injury' (*blaptesuai*) and since physically defending oneself certainly seems to entail injuring the attacker, physical defense would be a harm. Socrates, then, being opposed to harming another person, would be opposed to physically defending oneself.

Another reason, perhaps, is that as Socrates predicted very few people are on record as sharing Socrates principles of non-retaliation. The two most noted, however, are Gandhi and Jesus.¹⁸ Gandhi and Jesus, it is commonly supposed, were against any use of physical force.¹⁹ So, if we imagine that Socrates' position is similar to the positions of Gandhi and Jesus, Socrates would most likely be against the use of physical force as well, even to defend oneself.

Problems for the Jainist Interpretation

There are, however, some significant problems with supposing that Socrates' principles of non-retaliation are Jainist in nature. First, if Socrates were holding the claim that one ought *never* use any physical force against another person, his actions seem to prove otherwise. It is well-noted in the Platonic corpus that Socrates was in the military as a footsoldier, distinguishing himself in the battles of Potidaea and Delium. In fact, in the Symposium, Alcibiades claims that during the retreat from Delium, '[Socrates] made it plain from quite a distance away that if one tackled him, he would defend himself vigorously'

¹⁶As presented, the Jainist passage says nothing about injustice, but we might assume that the suggestion that one "should not" injure is a dictate of justice.

¹⁷The "ought" here I take to be a prudential ought--that is, if I want what is good for me I *ought* to do action X-- and not the "ought" of moral obligation. For reasons to think that the idea of moral obligation is absent from the Crito. See Penner, "Is the Crito a Treatment of Political Obligation"(unpublished).

¹⁸Martin Luther King, a disciple of Gandhian non-violence, may also be put in the category of noted supporter of the Jainist version of non-retaliation.

¹⁹Although Gandhi is commonly supposed to be against any use of violence whatsoever, some of his writings seem to contradict such a supposition. He writes:

I do believe that, where there is only a choice between cowardice and violence, I would advise violence. I would rather have India resort to arms in order to defend her honor than that she should, in a cowardly manner, become or remain a helpless witness to her own dishonor. (YI 11-8-20).

Nevertheless, because Gandhi is commonly thought to be a strict disciple of non-violence, for the purpose of this paper, I will interpret him to be strictly non-violent.

(221b). So, it does not seem that, if Socrates' principles are to be consistent with his actions, he cannot claim that one ought never use physical force to defend oneself.²⁰

A more formal problem is that the argument in support of the Jainist interpretation of 'never return harm for harm' contains two different conceptions of harm: For, we have seen that Socrates' claim in premise 5 that 'no one wants to be harmed' -- if it is to make any sense -- must contain an implicit comparison of the various options to determine which fail to maximize one's happiness.

A comparison of outcomes, however, seems absent from the use of 'harm' in 'defending myself harms the attacker' (premise 2). Instead, the premise seems to make harming the attacker an unconditional harm. Yet given the Socratic notion of harm we have been developing, we can only say that one is *harming* the attacker if in comparing all the other options, I am failing to maximize the attacker's happiness over a complete life.

So, if we incorporate into all the premises of the argument the notion that a harm involves comparing various options for a person's overall happiness, we get something like the following:

1. Since the psyche is more important than the body, ruining the psyche (by doing injustice) is more harmful than ruining the body (Crito.47e)
- 1* Harm = failing to maximize a person's happiness over a complete life from where s/he is now.
2. Defending myself fails to maximize the attacker's happiness
3. There is no difference between doing an injustice and doing a harm (Crito 49c)
4. Failing to maximize the attacker's happiness ruins my psyche (Crito.47e-48a) while not defending myself ruins my body (I get stabbed).
5. Since no one wants to fail to have his/her happiness maximized, I should not harm the attacker, preserve my psyche, and if necessary, sacrifice my body.

But given that the psyche is more important than the body (premise 1), am I failing to maximize the attacker's happiness by not defending myself, thus allowing him to succeed in an unjust act (premise 2)? If it is worse to ruin one's psyche than to ruin one's body, then, by allowing the attacker to perpetuate the harmful/unjust attack when I might prevent it, I am allowing the attacker to ruin his/her psyche and failing to maximize his/her happiness. Thus, premise 2 is false and the argument is unsound.

The Jainist interpretation of Socrates' principles of non-retaliation, then, does not conform to the facts about Socrates' life or his notion about what constitutes a harm. For if Socrates were to endorse the Jainist interpretation, he would be compelled to repudiate his earlier service in the Athenian army, which he never does. Moreover, the Jainist interpretation appears to deny the Socratic account of harm by assuming that in all circumstances one harms the attacker. Yet, if the attacker is facing the ruination of his psyche by committing an injustice, it is not at all clear that preventing him from doing so, even at the expense of his body, is harming him.

The Just Force Interpretation

The Just Force interpretation picks up where the Jainist interpretation runs into problems: The Just Force interpretation claims that when physically defending oneself (even when one physically damages the attacker) one is not, in fact, harming the attacker, so long as the attacker is trying to do an injustice/harm. Thus, the identity between harming and doing injustice is preserved in the self-defense case: one is doing nothing unjust by physically defending oneself, nor is one harming the attacker.

²⁰Two objections could be made against my claim that Socrates' actions would be inconsistent with his principles if he used physical force in war. (i) It could be argued that Socrates never actually had to fight in order to distinguish himself at the battle; he might have distinguished himself (e.g. saved others lives) without fighting. Though this option is possible, it is too implausible: I find it hard to believe that Laches would praise Socrates --if other soldiers behaved as Socrates, the defeat at Delium would have never occurred (La. 181b)-- if Socrates never struck a blow against the enemy.

(ii) Socrates could have developed his belief that one ought not use physical force *after* his term in the military was over. Once again, I think this unlikely. Not only does Socrates claim (Crito 49e) that he has held his principles of non-retaliation for a long time, but Alcibiades draws a picture of Socrates as a thoroughly ethical man *before* they served at Delium (Symp. 216bff).

But, one might say, how can breaking an arm of the attacker not be harming him? To answer this, we must remember two things:

(i) Though it is likely that, given the opportunity, Socrates would much prefer to convince the attacker that his attack is unjust,²¹ given the circumstances of the self-defense case, there are only two options, one must either defend oneself or suffer the unjust attack.

(ii) For Socrates, harm consists in comparing the various available options; those options which fail to maximize a person's happiness over a complete life are 'harms'.

Since Socrates believes that the psyche is more precious than the body, ruining one's psyche by doing an injustice is worse than ruining the body. So, if it comes down to whether one should ruin one's body by having an arm broken or ruin one's psyche by doing an injustice, the harmful option would be not to suffer the broken arm and do the injustice.

Given the restricted options surrounding the attack, then, Socrates' theory of harm suggests something quite outrageous-looking: if the attack is unjust, then physically wounding the attacker, perhaps even killing him, is not a harm at all. In fact, even if the defender kills the attacker, and thereby prevents the attacker from living with a corrupt, unjust psyche, in these circumstances, the defender is *benefiting* the attacker. For Socrates believes that life is not worth living with a corrupt psyche (Crito 47eff). But is there any reason to believe that Socrates would actually endorse this suggestion? I believe so. We might, as it were, modify the argument presented in the Jainist interpretation, focusing on the potential harm the attacker does to himself, rather than on the harm I might do to my psyche by defending myself.

1. Since the psyche is more important than the body, ruining the psyche (by doing injustice) is more harmful than ruining the body (Crito.47e)
- 1* Harm = failing to maximize a person's happiness over a complete life from where s/he is now.
2. By succeeding in the unjust attack, the attacker is ruining his psyche .
3. By failing in his attack when I physically defend myself, the attacker has his body ruined.
4. The attacker would not have his happiness maximized by succeeding in his attack.
5. I do not want to harm the attacker (i.e. fail to maximize the attacker's happiness)
6. If it is in my power, I should preserve the attacker's psyche , and if necessary, sacrifice his body by physically defending myself.²²

An analogy might serve to dispel some of the counter-intuitive features of the Just Force interpretation. Consider doctors doing triage in war: they do not have time for discussion with the patient; rather they must decide on their own what is in the best interest of the patient's health. They decide if it is in the patient's best interest to be without a limb, unable to walk, etc.. But in all these cases the triage doctor is only comparing different bodily states to achieve the patient's physical health.²³ As Socrates conceives harm, one must compare different states of the psyche in addition to the body, (bearing in mind that the psyche is more important than the body). Thus in our self-defense case, we must perform a bit of 'Socratic triage', comparing the worth of the attacker's body against that of his psyche. So, just as the triage doctor may have to sever a limb to do what's best for the patient's physical health, the defender may have to break a limb (or even kill) to do what's best for the attacker's overall happiness.

²¹I believe that, in general, Socrates would much rather try to convince people that what they are attempting to do is unjust (as Socrates does with Crito in the Crito 46bff), but in this case there is no time for discussion.

²²Socrates' concern that one should maximize other people's happiness when one has the power to do so may remind us of Socrates' self-appointed mission to Athens. Socrates believes that, by acting as a "gadfly to Athens", he is maximizing the happiness of the people he constantly questions.

...I went instead, to each one of you privately to do him, as I say, the greatest of benefits, and tried to persuade not to think of his affairs until he had thought of himself and tries to make himself as good and as wise as possible...(A.p.36c)

²³The doctor are only in a position to decide about a patient's physical health. They cannot judge whether a life is worth living or not. See L.a. 195bff.

We see further that the way in which we have been supposing Socrates distinguished himself as a footsoldier on the battlefield supports the Just Force interpretation. Given Socrates' account of human nature in the Protagoras--'No one who either knows or believes that there is another course of action better than the one he is following will ever continue on his present course when he might choose the better'--Socrates must have thought that going into battle was the best option open to him and would maximize his happiness. In addition, since Socrates believes that doing injustice ruins one's psyche and life is not worth living with a corrupt psyche (even less so than with a corrupt body), if he decided to fight Athens' enemies on the battlefield, he must have thought that he was not doing an injustice to members of the opposing army.²⁴

But we are still left with the question whether Socrates believed that he was doing no *harm* to the enemy when using physical force. The easy answer is, of course, that Socrates identifies doing injustice with doing harm. So, if he believes he does nothing unjust on the battlefield, it follows that he believes he does no harm either. But if we are to defend Socrates' identification of harm and injustice, we must find some evidence that he believed he was doing no harm on the battlefield.

The reason that Socrates did no harm on the battlefield might follow quite naturally from the fact that, according to his own theory, he must have believed that he was doing no injustice by fighting. He might then conclude that since he is doing no injustice, those on the opposite side of the war are doing something unjust. If the enemies were attempting an injustice, then, just as the person who uses physical force to defend him/herself, Socrates would, by his account of harm, be saving the enemies' psyches by preventing them from doing injustice. Thus, Socrates would not be harming the enemies at all; in fact he would be benefiting them by preventing them from doing injustice.

Problems for the Just Force interpretation:

Looking at the Just Force interpretation and Socrates' apparent endorsement of it, a modern philosopher might object that Socrates is overlooking the autonomy of the attacker. Is it our job to decide *for the attacker* that injustice is not in his best interest, and thus physically prevent him from attacking? After all, by Socrates' own theory, the attacker must be doing what he believes is the best option open to him at the time. Shouldn't we then allow him to make his own mistakes? Aren't we being rather paternalistic, deciding for the attacker what is a harm *to him*?

To begin, Socrates shows little interest in individual autonomy. Socrates is not opposed to paternalism, even against people's expressed wishes, so long as the paternalism achieves the maximal good. In the Lysis, Socrates states:

If [a king's] son had something the matter with his eyes, would [the king] allow him to touch them himself, if [the king] thought him ignorant of the healing art, or rather hinder him?--Hinder him--With regard to matters...into which we have acquired no insight, no one will ever allow us to act as we think proper, but all persons to the best of their power will hinder us from meddling with them....(Lysis 209e-210b)²⁵

If a person, through ignorance, will harm him/herself, Socrates sees nothing wrong in paternalistically forcing them to do otherwise. He makes no claim that it would be better to allow the ignorant person to harm themselves, in order that they may learn a lesson for the future.²⁶

²⁴One could suggest a rival interpretation based on certain authoritarian interpretations of the Crito: Socrates did think that physically engaging the oncoming enemy was an injustice, but because he felt a stronger obligation to do as the state commanded, he did what he believed to be the lesser injustice. But it is clear from the Apology that Socrates believes that the dictates of the state do not take precedence over the dictates of justice. When commanded by the Thirty tyrants to arrest Leon, Socrates disobeys, fails to give the tyrants any explanation for his refusal, and simply goes home. He tells the jury, however, that he did not arrest Leon because he believed it to be an unjust act and that "it mattered all the world to me that I should do nothing unjust or unholy"(32d). So, it seems that Socrates' concern is not ultimately what the state commands him to do, but not to do any injustice.

²⁵ "Lysis" in Plato: Collected Dialogues, J. Wright, trans. pp.145-168.

²⁶I would venture to guess that Socrates would find using harm to "teach a lesson" far too behaviorist and not intellectual enough to achieve its purpose. One should not cause or allow harm to happen to X in the hope that X draws the correct lesson from the incident. Instead one should prevent X from harming

Socrates paternalism is equally applicable to the self-defense scenario. We have been assuming that the attacker is in fact committing an injustice, trying to harm the defender. Thus, as the defender, if we know that the attack is unjust and that the attacker is ruining his psyche by perpetuating the attack, then we are entitled by Socratic principles to act paternalistically and to use 'the best of our power' to prevent the attacker from harming himself.²⁷

The Just Force interpretation, therefore, need not take into account the attacker's personal autonomy. The Lysis passage shows that Socrates is interested in what is *really best* for the king's son, even if that differs from what the son believes to be best. Much like a doctor doing triage, Socrates believes that one should do actions for a patient's best interest, even when those actions run contrary to the apparent wishes of the patient.

There is a price to be paid for accepting the Just Force interpretation, however. If Socrates defends himself in the Athenian army because he believes he is doing nothing unjust, it seems that his belief may be incorrect. Given Thucydides account of origin the Peloponnesian War, there is some good evidence that the war was caused, in part, by Athens perpetuating aggression on her neighbors, thus breaking a peace treaty with Sparta.²⁸

If Athens was guilty of unjustly causing the war, and if Socrates fought as a footsoldier on the unjust side of the war, Socrates' behavior might be explained in two ways:

- 1) Socrates incorrectly believed that Athens was on the just side of the war.
- or
- 2) Socrates took part in a war which he believed to be unjust.

Both options suggest a sort of failing on Socrates' part; the first suggests that Socrates was fairly naive about the political dealings that Athens had with her neighbors. The second suggests that Socrates took part in a war which he believed to be unjust, even though he claims never to have willingly done an injustice. If we must choose between these two failings, I would argue that we should accept the first. It is much easier to suppose that Socrates knew little of Athens' dealings with her neighbors than to suppose he would contradict his conviction that he would never willingly do an injustice. After all, Socrates states in the Apology (32d) that his divine guide forbade him to take part in politics, and he must fight for justice as a private citizen, not in public life.²⁹

Conclusion

One might be tempted to reject the Just Force interpretation because it explains Socrates two principles of non-retaliation-- never return injustice for injustice and never return harm for harm-- in a counter-intuitive manner. But I hope I have shown that other ways of interpreting Socrates doctrines of non-retaliation are even more counter-intuitive.

herself and then make X understand why the act was prevented. See Socrates rejection of punishment in the Apology.

²⁷Of course, with an actual case of self-defense, we cannot be so confident that we know that the attack is unjust and the attacker is trying to harm us. But even then, by Socratic principles, if the defender must act in the way that s/he believes to be the best, and additionally, the defender believes that (i) the attack is unjust and (ii) the attacker is harming both himself and the defender by attacking unjustly, the defender must, as a matter of psychological fact, try to prevent the attack if s/he believes defending oneself is best in these circumstances.

²⁸See Thucydides, History of the Peloponnesian War, Book I.

²⁹This same apolitical attitude can be seen in Socrates' reaction to the Thirty tyrants' command that he, along with four others, arrest Leon of Salamis. Socrates does not attempt to convince the tyrants of the injustice of arresting Leon nor does he try to convince the other four people to disobey the tyrants and not arrest Leon. Instead he simply dismisses himself from the entire situation, and goes home. Socrates will personally avoid politics, even if that leads to Leon being arrested by the other four citizens.

The Just Force interpretation, despite its counter-intuitive appearance, has two overwhelming advantages. First, it answers the alleged counter-example to Socrates' claim that there is no difference between doing an injustice to someone and harming him/her. By suggesting that a person using physical force is actually preventing the attacker from having a more miserable life, the Just Force interpretation shows that the defender neither does an injustice nor harms the attacker.

Second, it resolves the alleged inconsistency between Socrates' life and his philosophical positions. Socrates can, by interpreting his 'never harm others' principle appropriately, take part in Athenian battles at the same time believing that one ought never harm another person. For if Athens' enemies are risking the ruination of their psyches by attacking unjustly, then Socrates would not be harming them by wounding their bodies in route to stopping their injustice. In fact, according to the Socratic notion of harm and benefit, Socrates would be benefiting his enemies by saving their psyches from ruination.

Bibliography

Ayaro, translated by Muni Mahendra Kumar, in Understanding Non-Western Philosophy, Daniel Bonevac and Stephen Phillips eds., Mayfield Publishing, 1993.

Brickhouse, Thomas, and Nicholas Smith, Plato's Socrates, Oxford Univ. Press, 1994.

Koller, John M., and Patricia Koller, eds., A Sourcebook in Asian Philosophy, Macmillan.

Kraut, Richard, Socrates and the State, Princeton, 1983.

Penner, Terry, 'Is the Crito a Treatment of Political Obligation?' (unpublished).

Plato, Collected Dialogues, Edith Hamilton and Huntington Cairns, eds., Princeton Univ. Press, 1961.

----- 'Crito', Benjamin Jowett trans., in The Dialogues of Plato, Erich Segal, ed., Bantam Books.

----- Apology and Crito (Greek text and commentary), Louis Dyer, ed., Aristide D. Caratzas Publ., 1988.

Thucydides, History of the Peloponnesian War, Rex Warner, trans., Penguin Classics, 1972.