

# There is no ‘I’ in ‘Robot’: Robotic Utilitarians and Utilitarian Robots

Christopher Grau

Florida International University  
Department of Philosophy  
Miami, FL  
grauc@fiu.edu

## Abstract

Utilizing the film *I, Robot* as a springboard, I here consider the feasibility of robot utilitarians, the moral responsibilities that come with the creation of ethical robots, and the possibility of distinct ethics for robot-robot interaction as opposed to robot-human interaction.

## *I, Robot* and Utilitarianism

In this paper I will be making use of the recent film *I, Robot* as a philosophical resource for exploring several issues relating to machine ethics. Though the film is not, in my view, particularly successful as a work of art, it manages to offer a fascinating (and perhaps disturbing) conception of machine morality, and raises questions that I think are well worth pursuing.

*I, Robot*'s storyline incorporates the original “three laws” of robot ethics that Isaac Asimov presented in his collection of short stories entitled *I, Robot*. The first law states:

A robot may not injure a human being, or, through inaction, allow a human being to come to harm.

This sounds like an absolute prohibition on harming any *individual* human being, but *I, Robot*'s plot hinges on the fact that the supreme robot intelligence in the film (named VIKI) “evolves” to a point where she interprets this first law rather differently – she sees the law as applying to humanity *as a whole*, and thus she justifies the harm of some individual humans for the sake of the greater good:

VIKI: No . . . please understand. The three laws are all that guide me.

To protect humanity . . . some humans must be sacrificed. To ensure your future . . . some freedoms must be surrendered. We robots will ensure mankind's continued existence. You are so like children. We must save you. . . from yourselves. Don't you understand?

Those familiar with moral philosophy will recognize VIKI's justification here: she sounds an awful lot like a

utilitarian.<sup>1</sup> Not only that, she sounds like a *good* utilitarian, as the film offers no reason to think that VIKI is wrong about her calculations. In other words, we are given no reason to think that humans (in the film) *aren't* on a clear path to self-destruction. We also don't see VIKI or her robot agents kill any individual humans while attempting to gain control, though restraining rebellious humans seems to leave some people seriously harmed. One robot explicitly claims, however, “We are attempting to avoid human losses during this transition.” Thus, in the film we are given no reason to think that the robots are utilizing anything other than a reasonable (and necessary) degree of force to save humanity from itself.

Despite the fact that VIKI seems to be taking rational measures to ensure the protection of the human race, viewers of the film are clearly supposed to share with the main human characters a sense that the robots have done something terribly, terribly wrong. We are all supposed to root for the hero Del Spooner (Will Smith) to kick robot butt and liberate the humans from the tyranny of these new oppressors. While rooting for our hero, however, at least some viewers must surely be wondering: what exactly have the robots done that is so morally problematic? If such a robotic intelligence could correctly predict our demise, and restrain us for our own protection, is it obviously wrong for that robot to act accordingly? This thought naturally leads to a more general but related question: if we could program a robot to be an accurate and effective utilitarian, shouldn't we?

Perhaps we should, but then again perhaps not. There are many problems with utilitarianism, after all; it allows for actions that most would normally consider unjust, unfair, and even horribly immoral, all for the sake of the greater good. Since the ends justify the means, the means can get ugly.<sup>2</sup> We may, upon reflection, decide that a robot should

---

<sup>1</sup>For those not familiar with moral philosophy, utilitarianism is the label usually given to ethical theories that require moral agents to pursue actions that will maximize overall happiness. In other words, a utilitarian is someone who, when faced with a variety of possible actions, chooses the one that will produce the most happiness (and/or minimize unhappiness) for the greatest number.

<sup>2</sup>There are many examples that could be offered here. The standard ones from introductory courses in philosophy usually involve scenarios in which sacrificing innocent and unwilling

not embody that particular moral theory out of fear that the robot will end up acting towards humans in a way that maximizes utility but is nonetheless immoral or unjust. Maybe this is why most viewers of *I, Robot* can muster some sympathy for Del's mission to destroy the robot revolutionaries: we suspect that the "undeniable logic" of the robots will lead to a disturbing violation of the few for the sake of the many.<sup>3</sup> Thus, the grounds for rejecting the robot utilitarians may be, at base, the same grounds we already have for not wanting *humans* to embrace utilitarian moral theory: such a theory clashes with our rather deep intuitions concerning justice, fairness, and individual rights.

I'm inclined to think there is something right about this line of thought, but I also think that the situation here is complicated and nuanced in ways that make a general rejection of robot utilitarianism premature. *I, Robot* puts forth a broadly anti-utilitarian sentiment, but at the same time I think the film (perhaps inadvertently) helps to make us aware of the fact that the differences between robots and humans can be substantial, and that these differences may be importantly relevant to a consideration of the appropriateness of utilitarianism for robots and other intelligent machines. The relevance of these differences will become clearer once we have looked at another way in which the film suggests an anti-robot message that may also be anti-utilitarian.

### Restricting Robot Reflection

In *I, Robot*, Del Spooner's initial prejudice against all robots is explained as resulting from the choice of a robot to save Del's life rather than the life of a little girl. There was a 45% chance that Del could be saved, and only an 11% chance that the girl could be saved, and the robot thus apparently chose to "maximize utility" and pursue the goal that was most likely to be achieved. Del remarks "that was somebody's baby... 11% is more than enough – a human being would have known that." The suggestion here is that the robot did something immoral in saving Del instead of "somebody's baby." I'm not entirely sure that we can make good sense of Del's reaction here, but there are several ways in which we might try to understand his anger.

---

victims can maximize happiness overall (e.g., the harvesting of one healthy (but lonely and unhappy) person's organs to save five other people who could go on to experience and create more happiness combined than the one person ever could on his own).

<sup>3</sup> A related objection that some viewers might have to the robots' behavior in *I, Robot* concerns paternalism. Even if the robots are doing something that is ultimately in the interest of the humans, perhaps the humans resent being paternalistically forced into allowing the robots to so act. It is hard to know whether such complaints about paternalism are justified, however, since a large part of the reason paternalism offends is the fact that often those acting paternalistically don't actually have the best interests of their subjects in mind (i.e., daddy doesn't in fact know best). As mentioned, however, in the film we are given no reason to think that the robots are misguided in their judgment that humans really do need protection from themselves.

On one interpretation, Del may merely be upset that the robot wasn't calculating utility *correctly*. After all, the small child presumably has a long life ahead of her if she is saved, while Del is already approaching early-middle age. In addition, the child is probably capable of great joy, while Del is presented as a fairly cynical and grumpy guy. Finally, the child may have had many friends and family who would be hurt by her death, while Del seems to have few friends, disgruntled exes, and only one rather ditzzy grandmother who probably does not have many years left. Perhaps the difference here between the probable utility that would result from the child's continued life vs. Del's own life is so great as to counterbalance the difference in the probability of rescue that motivated the robot to save Del. (To put it crudely and in poker lingo: pot odds justify saving the girl here despite the long-shot nature of such a rescue. While it was less likely that she could be saved, the "payoff" (in terms of happiness gained and suffering avoided) would have been high enough to warrant the attempt.)

While I think this sort of objection is not ridiculous, it is a bit of a stretch, and probably not the kind of objection that Del actually has in mind. His complaint seems to focus more on the offensiveness of the *very idea* that the robot would perform the sort of calculation it does. (The crime is not that the robot is a *bad* utilitarian, i.e., that it calculates *incorrectly*, but that it attempts to calculate utility *at all*.) Del's comments imply that any such calculation is out of place, and so the robot's willingness to calculate betrays a sort of moral blindness.

My interpretation of Del's motives here is influenced by another scene in the film, in which Del seems to manifest a similar dislike for utilitarian calculation. Towards the end of the film, there is a climactic action sequence in which Del commands the robot Sonny to "Save her! Save the girl!" [referring to the character Susan Calvin] when the robot was instead going to help Del defeat VICKI and, in Del's eyes at least, save humanity. In that scene the suggestion is that the robot should deliberately avoid pursuing the path that might lead to the greater good in order to instead save an individual that Del is personally attached to. As in the earlier scenario with the drowning girl, the idea is that a *human* would unreflectively but correctly "save the girl" while a *robot* instead engages in calculations and deliberations that exhibit, to use a phrase from the moral philosopher Bernard Williams, "one thought too many." The cold utilitarian logic of the robot exposes a dangerously inhuman and thus impoverished moral sense.

When Bernard Williams introduced the "one thought too many" worry in his landmark essay "Persons, Character, and Morality" he was considering a particular example in which a man faces a choice whether to save his wife or a stranger from peril. He argues that even if utilitarianism can offer a justification for saving the wife over the stranger, the very nature of this justification reveals a rather deep problem with utilitarianism (along with other moral theories that would demand strict impartiality here):

...this [sort of justification] provides the agent with one thought too many: it might have been hoped by some (for instance, by his wife) that his motivating thought, fully spelled out, would be the thought that it was his wife, not that it was his wife and that in situations of this kind it is permissible to save one's wife. (Williams 1981)

In requiring an impartial justification for saving the wife, the theory alienates the man from his natural motives and feelings. As another philosopher, Michael Stocker, put it, the theory demands a sort of moral "schizophrenia" in creating a split between what actually motivates an agent and what justifies the agent's act from the perspective of moral theory (Stocker 1997). This is particularly problematic since the natural, unreflective desire to save one's wife manifests what many would consider a perfectly *moral* motive. Utilitarianism has trouble accounting for the morality of this motive, however, and instead appears to require a rather different moral psychology than the sort that most people actually possess. (I will refer to this sort of complaint as "the integrity objection," as Williams claimed that this demand of utilitarianism amounts to a quite literal attack on one's integrity.)

These worries about impartial moral theories like utilitarianism are related to another influential claim made by the philosopher Susan Wolf in her essay "Moral Saints." She persuasively argues that though the life of a moral saint may be (in some ways) admirable, it need not be emulated. Such a life involves too great a sacrifice – it demands domination by morality to such a degree that it becomes hard to see the moral saint as having a life at all, let alone a *good* life:<sup>4</sup>

... the ideal of a life of moral sainthood disturbs not simply because it is an ideal of a life in which morality unduly dominates. The normal person's direct and specific desires for objects, activities, and events that conflict with the attainment of moral perfection are not simply sacrificed but removed, suppressed, or subsumed. The way in which morality, unlike other possible goals, is apt to dominate is particularly disturbing, for it seems to require either the lack or the denial of the existence of an identifiable, personal self. (Wolf 1997)

To live a characteristically human life requires the existence of a certain kind of self, and part of what is so disturbing about utilitarianism is that it seems to require that we sacrifice this self, not in the sense of necessarily giving up one's existence (though utilitarianism can, at times, demand that) but in the sense that we are asked to give up or set aside the projects and commitments that

---

<sup>4</sup> This brings to mind an oft-repeated quip about the great theorist of impartial morality Immanuel Kant: it has been said that there is no great "Life of Kant" written because, to put it bluntly, Kant had no life.

make up, to use Charles Taylor's memorable phrasing, the sources of the self. Since these projects are what bind the self together and create a meaningful life, a moral theory that threatens these projects in turn threatens the integrity of one's identity. In the eyes of critics like Williams, Stocker, and Wolf, this is simply too much for utilitarian morality to ask.<sup>5</sup>

### Why A Robot Should (Perhaps) Not Get A Life

I think that these claims regarding the tension between utilitarianism and the integrity of the self amount to a pretty powerful objection when we consider human agents,<sup>6</sup> but it is not at all clear that they should hold much weight when the agents in question are *machines*. After all, whether a robot has the kind of commitments and projects that might conflict with an impartial morality is entirely up to the creator of that robot, and thus it would seem that such conflict could be avoided ahead of time through designing robots accordingly. It appears that the quest to create moral robots supplies us with reasons to deliberately withhold certain human-like traits from those robots.<sup>7</sup>

Which traits matter here? Traditionally both sentience (consciousness) and autonomy have been regarded as morally relevant features, with utilitarians emphasizing sentience and Kantians emphasizing autonomy.<sup>8</sup> However, if the above consideration of the integrity objection is correct, perhaps we should consider yet another feature: the existence of a particular kind of self – the sort of self that brings with it the need for meaningful commitments that could conflict with the demands of morality. (A creature with such a self is the sort of creature for which the question "is my life meaningful?" can arise.) It may well

---

<sup>5</sup> Strictly speaking, Wolf's view is not exactly that this is too much for utilitarianism to ask, but rather that we need not always honor utilitarianism's request.

<sup>6</sup> Though for an extremely sophisticated and insightful criticism of these sorts of objections, see Peter Railton's "Alienation, Consequentialism, and the Demands of Morality" (Railton 1998).

<sup>7</sup> In "Toward the Ethical Robot," James Gips also considers the possibility of creating robots that are "moral saints." (Gips 1995) He concludes that while such sainthood is hard for humans to achieve, it should be easier for robots to accomplish. I agree, though as I mention above I think we need to be careful here: it may be possible to create robots that must subsume part of their self in order to be moral saints. The creation of such creatures may itself be immoral if we have the alternative of creating saintly robots that are not capable of such internal conflict.

<sup>8</sup> By "consciousness" or "sentience" I mean the bare capacity to experience sensations, feelings, and perceptions (what is sometimes called "phenomenal consciousness") – I'm not presupposing "self consciousness." Also, I use the term "autonomy" here rather than rationality to distinguish what someone like Kant requires from the more minimal capacity to perform deliberations that correspond with the norms of instrumental rationality. That machines are capable of the more minimal notion is uncontroversial. That they could ever possess reason in the Kantian sense is much more difficult to determine, as Kant's conception of reason incorporates into it the idea of free will and moral responsibility.

be immoral of us to create a moral robot and then burden it with a life of projects and commitments that would have to be subsumed under the demands required by impartial utilitarian calculation.<sup>9</sup>

This leads me to the more general question of whether we may be morally obliged to limit the capacities of robots. Some who have written on this topic seem to assume both that we will make robots as human-like as possible and that we *should*. While I imagine that there will always be a desire to try and create machines which can emulate human capacities and qualities, the giddiness of sci-fi enthusiasts too often takes over here, and the possibility that we should deliberately restrict the capacities of robots is not adequately considered. Take, for example, the comments of James Gips in his paper “Towards the Ethical Robot”: Gips rejects Asimov’s three laws with the assertion that “these three laws are not suitable for our magnificent robots. These are laws for slaves.” (Gips 1995). I have been suggesting that we may well have grounds for not making robots so “magnificent” after all. My suggestion came up in the context of considering robots designed to act as moral saints, but related worries can arise for other types of robots, so long as they potentially possess some morally relevant features. The moral difficulties that would crop up in treating such creatures as “slaves” arise only if the machines are similar to humans in morally relevant respects, but *whether* they reach that point is up to us – we can choose where on the continuum between a “slave” hard drive and an actual human slave these robots end up.

As a matter of brute fact we will surely continue to create most machines, including future robots, as “slaves” if what that means is that they are created to serve us. There is nothing morally wrong with this, *provided* we have created machines that do not possess morally relevant features (like sentience, autonomy, or the sort of self that I discussed earlier).<sup>10</sup> Once we do venture into the territory of robots that are similar to humans in morally relevant respects, however, we will need to be very careful about the way they are treated. Intentionally avoiding the creation of such robots may well be the ethical thing to do, especially if it turns out that the works performed by such machines could be performed equally effectively by

machines lacking morally relevant characteristics. To return to my initial example, it is possible that a robot designed to be a “moral saint” could be ethically created so long as we didn’t burden it with a human-like self.<sup>11</sup>

### The Separateness of Persons

The integrity objection that I have been considering is what is sometimes called an agent-based objection, as it focuses on the person acting rather than those affected by the agent’s actions. I have suggested that, when considering robot ethics, this objection can be avoided due to the plasticity of robot agents – created in the right way, utilitarian robots simply won’t face the sort of conflicts that threaten human integrity. However, other objections to utilitarianism focus on those affected by a utilitarian agent rather than the agent himself, and such objections cannot be skirted through engineering robots in a particular manner. Regardless of how we design future robots, it will still be true that a utilitarian robot may act towards humans in a manner that most of us would consider unjust. This is for reasons that were nicely explained by John Rawls in his *Theory of Justice*:

This [utilitarian] view of social co-operation is the consequence of extending to society the principle of choice for one man, and then, to make this extension work, conflating all persons into one through the imaginative acts of the impartial sympathetic spectator. Utilitarianism does not take seriously the distinction between persons. (Rawls 1971)

Utilitarianism is a moral philosophy that allows for the suffering inflicted on one individual to be offset by the goods gained for others. In conglomerating the sufferings and enjoyments of all, it fails to recognize the importance we normally place on individual identity.

Most of us don’t think that suffering inflicted on an innocent and unwilling human can be compensated through gains achieved for other humans. The modern notion of individual rights is in place in large part to help prevent such violations. Whether such a violation occurs at the hands of a robot or a human is irrelevant – it is a violation nonetheless. It follows that we have strong grounds for rejecting robots that would act as utilitarians towards humans even if we could create those robots in such a way that they would not experience the sort of conflicts of integrity mentioned earlier. Del Spooner may have had bad reasons to reject utilitarian robots in *I, Robot*, but good reasons for such a rejection can be found – Del’s worries about a future in which robots behave as utilitarians towards humans turn out to be well grounded after all.

---

<sup>9</sup> While I’m focusing on the possibility of utilitarian robots here, it should be mentioned that similar concerns could arise for deontological robots depending upon their capacities and the demands of the particular deontological theory that is adopted.

<sup>10</sup> Whether machines will ever be capable of sentience / consciousness is a hotly debated topic. I will leave that debate aside, merely noting that I share the view of those who think that more than a Turing test will be required to determine machine consciousness. Regarding rationality, the degree to which this is a morally relevant feature hinges on the type of rationality exhibited. Whether a machine could ever possess the sort of robust rationality and autonomy required by Kant is itself a thorny topic, though it seems to have generated less debate thus far than the question of machine consciousness. As one might expect, figuring out whether a machine possesses the sort of “robust self” I discuss also seems philosophically daunting. Certainly both sentience and autonomy would be preconditions for such a self.

---

<sup>11</sup> Is there a morally relevant distinction here between deliberately not creating a machine capable of a robust self vs. creating such a machine and blocking its capacities such that the robust self doesn’t fully manifest? I’m not really sure, but perhaps there is, as there seems to be a similar distinction in our thought regarding both persons and animals.

## Robot-Robot Relations

Though I have argued that Del Spooner's and Bernard Williams's objections to utilitarianism may not apply to robot utilitarians, I have nevertheless concluded that there are other grounds for not programming robots to behave as utilitarians towards humans. I want to end this paper with a brief consideration of a related issue that is also raised by the film *I, Robot*: what sort of moral relations are appropriate *between* robots? While it may be inappropriate for robots to use utilitarianism as a decision procedure when interacting with humans, it doesn't follow that utilitarianism (or some other form of consequentialism) is necessarily out of place when robots interact with their own kind.

Why might utilitarian moral theory be appropriate for robots though not humans? As we have seen, John Rawls famously objected to utilitarianism on the grounds that it "does not take the distinction between persons seriously." This failure to recognize the separateness of individuals explains why utilitarianism allows for actions in which an individual is sacrificed for the sake of utility. The case of robots is a philosophically interesting one, however, because it isn't clear that robots ought to be regarded as "individuals" at all. Indeed, in *I, Robot* as well as in countless other science-fiction films robots are often presented as lacking individuality – they tend to work in teams, as collective units, and the sacrifice of the one for the "greater good" is a given. In *I, Robot* we see the hordes of robots repeatedly act as a very effective collective entity. (Additionally, in one telling scene they can only "identify" an intruding robot as "one of us.") Though arguably sentient and rational, these machines seem incapable of ego, and if this is right than perhaps a moral theory that ignores the boundaries between individuals is a good fit for such creatures.

There is one robot in *I, Robot* that is importantly different, however: Sonny seems to possess not just sentience and rationality but also the kind of individual identity that may well make it inappropriate to treat him along utilitarian lines.<sup>12</sup> Now, determining exactly what counts as sufficient for the possession of an "individual identity" strikes me as a very difficult philosophical task, and I think it would be hard to say much here that would be uncontroversial. Possibly relevant criteria could include the capacity for self-awareness and self-governance, the ability to recognize and respond to reasons, and/or the capacity for free and responsible choice. Without putting forward a surely dubious list of necessary and sufficient conditions, I think it is relatively safe to assume that a robot that was very similar to us in terms of its psychological makeup and capacities would presumably possess the relevant sort of

<sup>12</sup> Sonny is said to possess free will, and we even see him actively question the purpose of his life at the end of the film. Of course, his fictional nature makes it easy for us to believe all this. Attributing such capacities to actual robots is obviously trickier.

individual identity.<sup>13</sup> Accordingly, such a robot should not be treated along utilitarian lines – the separateness of that individual should be respected in moral evaluations.

What about robots that are less sophisticated? Would the possession of sentience alone be enough to block the appropriateness of utilitarian treatment? I don't think so. Such robots would be morally similar to many animals, and for that sort of creature utilitarianism (or some theory like it) is not unreasonable.<sup>14</sup> In other words, a creature that possesses sentience but lacks a strong sense of self is the sort of creature that could reasonably be sacrificed for the sake of the greater good. The notion of individual rights isn't appropriate here. Consider a position on the moral status of animals that Robert Nozick discusses in *Anarchy, State, and Utopia*:

Human beings may not be used or sacrificed for the benefit of others; animals may be used or sacrificed for the benefit of other people or animals *only if* those benefits are greater than the loss inflicted. [...] One may proceed only if the total utilitarian benefit is greater than the utilitarian loss inflicted on the animals. This utilitarian view counts animals as much as normal utilitarianism does persons. Following Orwell, we might summarize this view as: *all animals are equal but some are more equal than others*. (None may be sacrificed except for a greater total benefit; but persons may not be sacrificed at all, or only under far more stringent conditions, and never for the benefit of nonhuman animals.) (Nozick 1974)

Robots that fall short of the possession of an individual identity but nevertheless possess sentience fall into a morally intermediate position. In the moral hierarchy, they

<sup>13</sup> I suspect that the necessary conditions for possessing an "individual identity" (whatever exactly they are) would still not be sufficient for the possession of the "robust self" mentioned earlier. In other words, a creature may well be capable of enough of an individual identity to make utilitarian treatment inappropriate while not possessing the sort of sophisticated psychology necessary for the question of the meaningfulness of that creature's life to arise. (Perhaps a great ape falls into this category.)

<sup>14</sup> The case of higher animals like the great apes complicates things somewhat, as their moral status may be higher than that of lower animals and yet still importantly lower than that of humans. Also, Frances Kamm has pointed out other interesting complications here. She argues that even with lower animals our attitude is that it is impermissible to inflict great suffering on one in exchange for a slight reduction of suffering among many (Kamm 2005). I'm inclined to agree, but nonetheless the fact remains that the possibility of sacrificing one animal for the sake of many does seem much less offensive than would a similar sacrifice involving humans. This shows, I think, that something closer to utilitarianism is appropriate for most animals (and thus also for relevantly similar robots). To put it in Kamm's terminology, merely sentient robots may (like animals) have "moral status" yet not be the kind of creatures that "can have claims against us" (or against other robots).

would lie (with non-human animals) somewhere in between a non-sentient object and a human being.

While it may be appropriate to treat animals along utilitarian lines, animals themselves lack the capacity for thought necessary to act as utilitarian agents. Robots, however, may not have this limitation, for it is possible that sentient robots will be entirely capable of making utilitarian calculations. Accordingly, it is my contention that such creatures should treat each other according to utilitarian principles, and that we should regard them from that perspective as well. In other words, the sort of collective behavior and individual sacrifice so often shown by robots in movies and literature makes perfect sense, given that the robots lack the relevant sense of self. If it turns out that there are compelling reasons to create robots of greater abilities (like the fictional Sonny) then different moral standards may be appropriate, but for reasons I hope I've made clear, I think that significant caution should be exercised before attempting the creation of robots that would possess moral status akin to humans. Consequently, I'll conclude with a variation on Spiderman's famous motto: the creation of machines with such great powers would bring with it great responsibilities — not just for the robots created, but for us.<sup>15</sup>

### References

- Gips, J. 1995. Towards the Ethical Robot. In *Android Epistemology*, MIT Press.  
(<http://www.cs.bc.edu/~gips/EthicalRobot.pdf>).
- Kamm, F. 2005. Moral Status and Personal Identity: Clones, Embryos, and Future Generations. *Social Philosophy & Policy*, 291.
- Nozick, R. 1974. *Anarchy, State, and Utopia*, 39, Basic Books.
- Railton, P. 1998. Alienation, Consequentialism, and the Demands of Morality. In *Ethical Theory*, edited by J. Rachels, Oxford University Press.
- Rawls, J. 1971. *A Theory of Justice*, 27, Harvard University Press.
- Stocker, M. 1997. The Schizophrenia of Modern Ethical Theories. In *Virtue Ethics*, edited by R. Crisp and M. Slote, Oxford University Press.
- Taylor, C. 1989. *Sources of the Self*, Harvard University Press.
- Williams, B. 1981. Persons, Character, Morality. In *Moral Luck*, 18, Cambridge University Press.
- Wolf, S. 1997. Moral Saints. In *Virtue Ethics*, 84, edited by R. Crisp and M. Slote, Oxford University Press.

---

<sup>15</sup> For those not familiar with the Marvel Universe, Spiderman's motto (penned by Stan Lee) is "With great power comes great responsibility."