Sophisticated Consequentialism and Psychological Coherence

In “Alienation, Consequentialism, and the Demands of Morality,” Peter Railton introduces an approach to consequentialism that has since become influential.\(^1\) His approach, which he calls ‘sophisticated consequentialism’, claims that a consequentialist ought to be motivated by whichever set of motivations will in fact lead to maximizing value, instead of being motivated to directly maximize value. As a result, a consequentialist might maximize value by having dispositions to commit acts that may not themselves maximize value. I concern myself here with the psychological tension that might exist in a person with such apparently conflicting commitments: how may a person that is, on one hand, committed to believing that consequences determine how people ought to act, on the other hand, be motivated to generally act in ways that pay no direct attention to consequences? In particular, I will argue that, on Railton’s picture, a consequentialist must pay attention to two distinct and inconsistent sets of reasons. First, it is difficult to see how a person can really make decisions in that way, and second, a person situated in the way Railton imagines is not actually in a position to maximize value.

But before I get to a more rigorous statement of what I will call the criticism from psychological incoherence, I need to lay the scene. The criticism from psychological incoherence is related to another objection to which Railton was directly responding, the alienation objection. I first set out this objection before moving to a statement of sophisticated consequentialism that handles the alienation objection. Once the first dialectic is on the table, I present the criticism from psychological incoherence.

The alienation objection claims that the modes of thinking about one’s own central projects that

are required of a consequentialist alienate consequentialists from those central personal projects. One of the major tenets of standard consequentialisms is that we must weigh the various impacts on various people from each of the relevant actions in the current circumstance and pick the action for which the calculations come out most positive. Such an approach requires that the consequentialist treat the projects of each person as considerations of a certain size weighing in a certain direction. In particular, even the most important projects of the agent herself around which the life of the agent has been constructed must be treated in this way. But to treat projects that are sufficiently central to one's life as just another consideration to be weighed, similar to many other considerations, is to lose grasp of the role they have played in the development of the agent's life so far. As Bernard Williams, a major proponent of this criticism, says, “It is absurd to demand of such a man, when the sums come in from the utility network which the projects of others have in part determined, that he should just step aside from his own project and decision and acknowledge the decision which utilitarian calculation requires. It is to alienate him in a real sense from his actions and the source of his action in his own convictions.”

Consider Railton's example of a deeply committed intimate relationship. John shows a deep and abiding concern for his wife Anne through his words and actions. When a friend notes the extraordinary character of John's concern for Anne, John responds by explaining that, really, it is no trouble at all; his emotional connection to Anne means that making sure she is happy makes him happy, and, moreover, because they are so close, he knows which things will make Anne the happiest, and the world is better off when people who are in that kind of position leverage it to make their loved ones happy. In this case, it seems like John is acting in a way that clearly embodies a consequentialist approach to morality, but at the same time is displaying a certain emotional distance from his

---

3 Williams, p. 49.
relationship with Anne that seems odd. We might expect that Anne would be unhappy that John's relationship with her has such an indirect effect on the way John makes decisions or that his motivations for acting on her behalf are always mediated by considerations of impartial value. We may even wonder whether someone such as John could ever really enter into an intimate relationship of the kind that plays such an important role in a satisfying life, since his putative emotional connection to Anne has so little impact on his decision-making. It seems that part of the significance of these relationships is that they impact our decision-making directly and strongly. If John lacks such a central marker of the significance of these relationships, perhaps we ought to be skeptical of the thought that the relationship plays the other significant roles we might expect it to play. Moreover, those relationships are two-way streets, and the people with whom he might enter into these relationships could be seriously deterred by his attitude.

If it is true that someone living by what seems to be such a paradigmatically consequentialist way of making decisions must thereby miss out on such a crucial part of the good as intimate personal relationships, that might be a reason to take consequentialism less seriously. As Railton puts it, “If we were to find that adopting a particular morality led to irreconcilable conflict with central types of human well-being—as cases akin to John's... have led some to suspect—then this surely would give us good reason to doubt its claims.”

**Railton's Sophisticated Consequentialism**

Railton's central insight is that, if Williams says that thinking like a consequentialist does not maximize value, then so be it; we need not think like consequentialists. We can instead think like

---

4Railton, p. 135-137.
5Railton, p. 139.
friends and family members and thereby actually maximize value. Consequentialism, as a claim about what makes actions right, need not be tied to any particular way of making decisions. Railton develops this line through the example of Juan and Linda that correlates to a sophisticated consequentialist version of the John and Anne case. When a friend comments on Juan's deep commitment to and concern for his wife, Juan responds that he loves his wife and that their relationship is so important and long-standing at this point that doing things for her just comes naturally. But this response puzzles the friend a little, given the kind of person Juan is; it is good that Juan and Linda have such a close relationship, but couldn't Juan help so many more people, people in much greater need than Linda, if Juan turned his attention and resources to helping them?\(^6\) Railton has Juan say in response:

“‘It's not easy to make things work in this world, and one of the best things that happens to people is to have a close relationship like ours. You'd make things worse in a hurry if you broke up those close relationships for the sake of some higher goal. Anyhow, I know that you can't always put family first. The world isn't such a wonderful place that it's OK just to retreat into your own little circle. But still, you need that little circle. People get burned out, or lose touch, if they try to save the world by themselves.’”\(^7\)

By taking this line, Railton provides a kind of answer to Williams' criticism from alienation. That argument turns on the agent having a certain attitude to her own major projects, including relationships, namely the attitude of having to weigh the importance of those projects against a variety of other considerations before acting. But Railton denies that a consequentialist really has to weigh anything in that way; only someone following the traditional utilitarian decision procedure must do that, and Railton's central insight is that a consequentialist need not follow that decision procedure precisely because doing so does not maximize value. An agent, per Railton, ought to decide, not as a traditional utilitarian would, but as a friend or partner would, because doing so prevents alienation and the bad consequences associated with alienation.

---

\(^6\) Railton, p. 150.

\(^7\) Railton, p. 150.
Psychological Incoherence

However, sophisticated consequentialism has its own problems that are still related to the issues Railton was attempting to solve. The criticism I offer comes originally from Michael Stocker, despite the fact that it predates Railton's writings. His claim is that if an agent follows a theory that requires that agent be motivated by different considerations than those the agent values, the agent is thereby robbed of one of life's major goods: psychological coherence. Stocker took his argument to apply to most or all contemporary ethical theories; whether most contemporary moral theories fall prey to this criticism or not, Railton's sophisticated consequentialism most certainly does.

The centerpiece of Railton's theory is the claim that sophisticated consequentialists need not be motivated by the reasons that they think, in the end, justify their actions. Stocker's point is that this situation poses a problem for the agent. There are two sets of considerations the agent thinks are important: there are the relationship-based considerations and the consequentialist considerations, and in relevant cases they are inconsistent. There is no good reason to think that these sets of considerations will tend to support the same action, and as a result, the sophisticated consequentialist faces two challenges. First, it is not clear how she is to decide between doing what she is motivated to do, as endorsed by the theory, and doing what the theory directly says to do. Second, the sophisticated consequentialist seems to be under an enormous amount of psychological stress from the need to adjudicate between her own motivations and the direct dictates of her moral theory.

Even if the agent gets lucky and both sets of considerations endorse the same action, there is still the question of which set to use, which set to actually consider when trying to determine what to do. If the answer is the consequentialist reasons, then it is no longer clear why the relationship reasons—

---

9Stocker moves between using the language of 'schizophrenia' and 'disharmony,' but I will use 'incoherence.' I intend no difference in meaning to follow this difference in terminology.
that ought to motivate, according to the moral theory—are important, or how it is that the agent avoids alienation. In any case, this is not the option Railton's view endorses.

Instead, Railton's innovation is that the sophisticated consequentialist need not pay any attention, in the moment, to consequentialist reasons. These reasons need not serve any motivating purpose at the moment of decision-making and, often, ignoring these reasons results in the sophisticated consequentialist bringing about the best consequences. But for Stocker, it is important to realize that the consequentialist reasons still have force for the agent. If we are to make sense of the sophisticated consequentialist's commitment to consequentialism, it had better mean something to the agent that one action or another maximizes value. Psychologically, it seems difficult to imagine an agent who is really committed to a particular criterion of rightness, but for whom considerations relating to that criterion do not enter into any decision of how to act. How is the sophisticated consequentialist to make sure that the new, non-consequentialist, relationship-based reasons are in fact still keeping in line with consequentialist ends?

Railton says that the sophisticated consequentialist should abide by a counterfactual condition: she would not act as she does if doing so were not compatible with leading a life that is morally defensible from a consequentialist perspective. But this gives no particularly good direction to the sophisticated consequentialist about how she is to accomplish this end. It seems the sophisticated consequentialist must have two sets of reasons: mostly, the sophisticated consequentialist abides by a common, everyday set of reasons, but she sometimes makes reference to a higher-order regulative set of reasons. How can the sophisticated consequentialist keep these sets of reasons apart? When ought the sophisticated consequentialist subject her life to consequentialist assessment? Attempting to integrate these competing sets of reasons into one life undoubtedly causes serious psychological stress:

---

10 Stocker argues that this option leads to a much impoverished life, lacking much of what makes life worth living. See p.460-461, among other places.
11 Railton, p. 151.
one must be committed to consequentialist aims while rarely actually considering consequentialist reasons. “It is bad enough to have a private personality, which you must hide from others; but imagine having a personality that you must hide from (the other parts of) yourself.”\textsuperscript{12} For Stocker, the sophisticated consequentialist believes that, in the end, consequentialist considerations are all that matter, but she must never act on that belief.

The pressure on the agent from the conflicting sets of reasons becomes even stronger when we consider the possibility that the sets of reasons might produce different answers about what to do. Given what has been said so far, there is no reason to believe a sophisticated consequentialist's relationship-based reasons must always align in the end with her consequentialist reasons. Indeed, we might plausibly think that by taking the good of our loved ones into account directly, we upset the delicate balance of assessment necessary to recommending actions that produce maximally valuable outcomes. If this does happen, with which set of reasons does the agent side? She can side with the assessment grounded in her deepest moral beliefs, or she can side with the assessment based on those people whose place in her life is most central. Perhaps that decision can be made, but the person who has to make it is diminished by the necessity.\textsuperscript{13}

A defender of Railton might respond that the agent need not know that the two sets of reasons conflict: the agent ought to abide by the relationship-based reasons, not the consequentialist reasons, and if the agent does not know, then there is no psychological turmoil. However, this response gives strength to the earlier description of the divided mind. Perhaps Stocker's quote about private personalities sounds over-dramatic; perhaps it was not entirely clear why the sets of reasons must be kept so harshly apart. Now, however, it seems that in order to prevent the internal struggle described

\textsuperscript{12}Stocker, p. 458.
\textsuperscript{13}It is interesting to note that the structure of the criticism from incoherence is exactly analogous to that of the criticism from alienation. Both argue that following a consequentialist moral theory prevents an agent from experiencing one of the major goods of life. The difference is that the major good that the criticism from alienation is concerned with is intimate relationships, whereas the criticism from incoherence is concerned with psychological unity.
above, the agent must be very careful to isolate the consequentialist reasons from the relationship-based reasons, lest they be seen to conflict. Stocker argues that such isolation creates a significant gap in the mental life of the agent. Such a sophisticated consequentialist might say: “Consequentialist reasons are the reasons that matter, because consequentialism is the true moral theory. But nevertheless I ought to pay no attention whatever to those consequentialist reasons, or they might truly tell me that I ought not to act as I am motivated to act...” Even so stringent a step may be insufficient. There still must be some consequentialist check that makes certain that the agent’s life abides by the counterfactual condition. This check may easily conflict with the relationship-based reasons and can produce internal conflict.¹⁴

Williams’ criticism from alienation and Stocker’s criticism from incoherence therefore form a dilemma for the consequentialist. Either she must act directly on consequentialist motivations and be alienated from her intimate relationships, or she takes Railton’s approach: she is motivated directly by her friends and is thereby made incoherent.

Railton does offer the example of Juan and Linda, a sophisticated consequentialist version of John and Anne, to make his account more plausible, but there is no clear picture of how Juan’s decision-making actually works. Railton puts in Juan’s mouth both the claims that family must not always come first and that responding to Linda’s needs and desires are “‘almost a part of [him],’”¹⁵ but we have no idea how it is that Juan balances those considerations. Indeed, it seems that if he does ever explicitly balance those considerations, he is on the path to alienation.

But once we’ve gotten clear on the criticism from incoherence and the dilemma it forms with

---
¹⁴A defender of Railton might take a more extreme step and say that this kind of internal conflict can be averted if the agent does not actually believe consequentialism. Such an agent would just have one set of reasons. On this picture, consequentialism would endorse that no one would be consequentialist; they would merely act as consequentialism would have them act, without believing that consequentialism is true. Consequentialism would be self-effacing. But on such a picture, there is no way to reliably guarantee that people will continue to act consistently with consequentialism, given that circumstances may change. Elinor Mason assesses the possibility that consequentialism could be self-effacing (and related issues) in “Do Consequentialists Have One Thought too Many?” Ethical Theory and Moral Practice 2: 243-261, 1999, p. 250-251.
¹⁵Railton, p. 150.
the criticism from alienation, a further problem arises. Even if we were to grant that there is some course of action which maximizes value by giving up value from some individual decisions and thereby gaining the value of some intimate relationships, how is it that a sophisticated consequentialist will find and follow that particular set of actions? Surely not every course of action that is consistent with following relationship-based reasons is better than every course of action that follows consequentialist reasons directly; it is possible to sacrifice too much value by paying too much attention to one's loved ones. As Railton's Juan says, “Anyhow, I know that you can't always put family first. The world isn't such a wonderful place that it's OK just to retreat into your own little circle.”\textsuperscript{16}

The problem is that it seems Juan does not have the right resources to determine what the correct balance is. He cannot tell what the right balance is by assessing all the options; he risks alienation by trying to weigh the full implications of his choices. But if he cannot pay attention to the consequentialist reasons because of alienation, how can he possibly find the right balance?

If it turns out that the only way to gain the value of close relationships is to not pay attention to consequentialist reasons, and the only way to maximize value involves the value of close relationships, but also that in order to figure out which precise course of action involving the value of close relationships maximizes value we must pay attention to consequentialist reasons, then Railton is in a bind. Either paying attention to consequentialist reasons prevents the agent from having the friendships that will allow the agent to maximize value, or not paying attention to consequentialist reasons prevents the agent from knowing what precise course of action will maximize value. Either the agent knows what to do to maximize value, or the agent can do it, but not both.

I do not take these arguments to be decisive against consequentialism. But they do show that consequentialists must be much more concerned with their own psychology and motivations than they

\textsuperscript{16}Railton, p. 150.
have traditionally been. Any defensible consequentialism that makes Railton's move of strongly
distinguishing decision procedure from criterion of rightness must come tied to a clear story of how it
is that the advocated decision procedure is psychologically consistent with accepting the
consequentialist criterion of rightness.