In this paper, I will argue that Kant provides us with a plausible account of morality. To show that, I will first offer a major criticism of Kantian moral theory, by explaining Bernard Williams’ charge against it. I will explore his understanding of the Kantian theory, and then explain what he finds objectionable about it. This criticism will make up the first part of the paper. In the second part, I will attempt to defend the Kantian theory by appealing to Christine Korsgaard’s alternative reading. This reading, I will argue, accommodates Williams’ worries and avoids the charge altogether. Finally, I will attempt to show that in avoiding Williams’ charge, this alternative reading does not lose sight of the greater task it claims to undertake, namely being a plausible account of morality.

Williams believes that there is a reading of Kantian moral theory that is characterized by an “Impartiality and … indifference to any particular relations to particular persons” (Williams 1981, 2). He further believes that this reading “Requires abstraction from particular circumstances and particular characteristics of … the agent, except in so far as these can be treated as universal features of any morally similar situation” (Williams 1981, 2). Exactly what it is in Kant’s moral theory that lends itself to such reading, Williams does not make clear. However, it is not awfully hard to speculate about such a feature.

After offering the first statement of the categorical imperative, Kant examines three examples that are aimed at demonstrating its practical application (Kant 1998, 4:422). In these three examples one can detect a pattern emerging. First, the agent has to state the decision she is deliberating about in the form of a maxim. Then, she must generalize the maxim in the form of universal law. Finally, she must decide whether or not it is possible to will such a law as a “law of nature” (Kant 1998, 4:422). Now this process clearly involves some sort of abstraction. Williams claims that this abstraction requires the agent to abstract from who she is. She must regard the situation from an impartial standpoint and make a decision that would be binding on any arbitrary rational being.

Williams further claims that this outcome is undesirable. To illustrate the point, Williams uses the following example (Williams 1981, 17): Imagine two people who are equally in danger. A man has the opportunity to help one of them. It seems obvious that the man has to be completely impartial in his decision about which person to help. But, there is more to the story: One of the people who are in danger is his wife. Now, Williams thinks that it is hard to doubt that the man should resolve to favor his wife over the stranger. Although a Kantian would arguably say that such a decision is permissible, Williams thinks that there still is a problem. He worries that, the Kantian theory would “Provide the agent with one thought too many” (Williams 1981, 18). Before clarifying what this phrase means, let us take a detour and define a central term in Williams’ moral discourse.

Williams considers one’s “projects” and “commitments” as reasons that, “Constitute conditions of there being … a future” for the agent (Williams 1981, 11). That is, projects and commitments are things that make one’s life worth living in the first place and thus are fundamental to one’s existence. Since the choice to continue living is the most fundamental choice one could make, it follows that any choice that one would make is subordinate to one’s commitments and desires.
Now we begin to see the problem that Williams aims to present. Since, as we have seen, the Kantian account requires the agent to engage in a sort of thought experiment to make any choice, it requires her to deliberate about her most dearly held commitments. In extreme cases, it could even require the agent to overlook her desires, because from the impartial viewpoint it would be impossible to will them. This is problematic. What is more, the less extreme case, in which there is no conflict between what morality requires and what one’s projects are, is no less problematic. That is, even if there were no conflict, as in our rescue example above, the agent would still have to abstract from her projects and assume an impartial view of the situation. In the above case, the Kantian theory would require the man to take a moment and think whether it is permissible to rescue his wife or not and to will such a decision as a universal law. With the central role that Williams attributes to one’s commitments, this seems objectionable. That is, Williams wants to say that the man shouldn’t save his wife because he can universalize a maxim that makes it permissible; rather he should save her because she’s his wife! Thus, Williams concludes, “Such things as deep attachments to other persons will express themselves in the world in ways which cannot at the same time embody the impartial view, and they also run the risk of offending against it” (Williams 1981, 18).

Up to this point, I have characterized a certain reading of the Kantian theory that Williams found objectionable. Further, I have shown what it is about this reading that Williams objects to. In what follows, I will use Korsgaard’s reading of the Kantian moral theory to show that a deeper understanding of it avoids the charge Williams makes altogether.

Before setting out to do that, though, I would like to note that Korsgaard, like Williams, justifiably holds that our projects and commitments, what she calls our identities, are at the center of morality. Your identity, she writes, is “A description under which you value yourself, a description under which you find your life to be worth living and your actions to be worth undertaking” (Korsgaard 1996, 101). As such, our identities are things that oblige us unconditionally (Korsgaard 1996, 102). Consider the following:

Suppose that I identify myself with being an A. This means that I take myself to be an instance of A. As such, I commit myself to doing anything that would constitute being an A. What gives being an A normative authority over me is the fact that I take myself to be an A. Now, if for some reason I fail to comply with an obligation that being an A requires, then I would not be an A (Korsgaard 1996, 102). I would fail myself in my determination to be what I chose to be, an A. And that is an undesirable thing: After all, what good is my commitment if I myself don’t find it valuable enough to want to keep it? Thus, Korsgaard and Williams justifiably agree on this much: Projects and commitments, or one’s identity in Korsgaard’s language, are the foundational motives for action. If there is anything that has normative force, it must be one’s commitments.

The question to ask, then, is this: How is Korsgaard’s agreement that one’s identity is central to any account of morality consistent with Kantian moral theory, if Williams’ understanding of the Kantian moral theory is correct? In responding to this question, Korsgaard offers an alternative reading of Kantian moral theory. I will attempt in the rest of this paper to explain what this reading is, and how it allows Korsgaard to, on the one hand, be committed to the Kantian moral theory and on the other, consistently hold that one’s commitments are the sort of things that give one reasons to continue living and thus are foundational.

Let us ask: What is it to will something (Korsgaard 1996, 93)?
Kant began his project in the *Groundwork* by examining the authority of our reflective thinking about the choices we face in our lives. He thought that this authority could be partly explained by imperatives that are aimed at some end. For example, I resolve to drink a glass of water, because I want to satiate my thirst. However, he realized that those ends themselves need to derive their authority from somewhere. Thus, he proposed there must be imperatives that derive their authority from themselves. As such, these fundamental principles cannot be aimed at yet another end themselves. The categorical imperative is absolute in the sense that it derives its authority from itself (Kant 1998, 4:414). Kant wrote,

“There is, therefore, only a single categorical imperative and it is this: Act only in accordance with that maxim through which you can at the same time will that it becomes a universal law” (Kant 1998, 4:421)

There is nothing in this statement that reveals what the content of the law is. All that this statement of the categorical imperative says is that the agent has to act according to some law (Korsgaard 1996, 98, 120). That is, if there is to be any decision, there has to be some reason. That is explained by the fact that human beings, the kind of agents that we are discussing when talking about morality, are reflective agents (Korsgaard 1996, 93). If that is true, we think and wonder about the situations we face, and in order to get ourselves to act we need some kind of principle on which to base our reflections. In other words, Kant tells us that the structure of our reflection is such that we need laws to act.

What is more, these laws cannot gain their authority from outside of the agent (Korsgaard 1996, 100). Even if a law exists outside of the agent, it has no force on the agent unless the agent has some representation of it in her. That is, unless the agent herself consents to a law, the law has no binding effect on her. Thus, the categorical imperative, if it is to have any force, must be represented in the agent. In other words, the categorical imperative is the law that the self formulates for itself.

What we have established is that there is a structure to the way we will things. When we face a decision to will something, we first ask ourselves a normative question. Then, we set out to answer the question according to the kind of laws we have. Central in this picture is that there is something special about this set of laws. They are strictly personal, or in the Kantian language, autonomous. They are not imposed on the self from the outside, and they are not formed based on the kind of ends that one might have. With this, we have established what these laws might look like. So, the question now is: What is the content of the categorical imperative? What is there that is wanted categorically?

According to Korsgaard the representation of the laws that the self requires in order to make judgments and resolutions are encapsulated in our practical identities. “It is the conceptions of ourselves that are most important to us that give rise to unconditional obligations” (Korsgaard 1996, 102). As we have seen, for Kant and Korsgaard our conceptions of ourselves and our identities are indeed personal imperatives. In other words, Korsgaard seems to point to what we were looking for, something that would fit the criteria we have defined above. First, one’s identity provides the self with norms and standards to judge and make decisions, and thus it satisfies the condition to be an imperative. Second, one’s identities are the representations that one adopts for their own sake and for no greater end. And third, one’s identity is not imposed
upon the self from the outside\textsuperscript{1} and thus it satisfies the condition to be personal. Our identities, then, are the will’s way of devising laws to govern itself.

Now that we know what Korsgaard’s account of Kantian moral theory is, we can see that it indeed avoids the problem Williams raised. Let us recall that Korsgaard agrees on two issues with Williams. First, she agrees that one’s projects and commitments, one’s conception of one’s self, is what gives one reason to continue living in the first place and thus she agrees that they have a central role in one’s decision-making processes. Second, she agrees with Williams that the Kantian theory involves some level of abstraction. By that, I mean the abstraction involved in stepping back from one’s situation and reflecting on one’s options. Indeed she thinks this abstraction just is what constitutes a human being: we are “essentially reflective” (Korsgaard 1996, 92). However, what separates her from Williams is this. She disagrees that this abstraction translates into the problem Williams points to, because she takes Kant to have positioned one’s identities rightly in the center of morality. Thus, although there is abstraction involved in one’s asking oneself the normative question, one still provides the answer to this question by appealing to one’s identity, projects, and commitments. Thus, it would be impossible for one to come up, in his Kantian reflective exercise, with a decision that is in conflict with one’s projects. That is because it is precisely one’s projects that provide the grounds for a response to the normative question. If this is how we reach our decisions, then how could one, as Williams claimed, abstract from one’s identities and still be able to act at all? In the case of the man who has to make a decision whether to favor his wife, it is clear, that he does not abstract from the person he is, because if he did abstract from his identity as a human being, then it would have been impossible for him to make a decision at all (Korsgaard 1996, 121). He would ask the normative question and then, nothing. He would be paralyzed, so to say. He acts because he has some principle in him that tells him to do one thing rather than another.

Thus, Korsgaard’s account of Kantian moral theory avoids the problem raised by Williams. What is now in order is a general examination of Korsgaard’s account to make sure that in responding to this charge it has not fallen short of its original aim of being a viable account of morality.

One worry that could be raised is this. If it is an agent’s mere endorsement of an identity, whatever it might be, that gives it its authority, then what prevents an agent from justifying obviously impermissible acts by simply endorsing an identity that would accommodate such acts? The answer to this question cannot be simply “nothing.” Obviously, there must be something wrong with a moral theory that would allow an individual to, say, torture innocent babies on the condition that the individual’s identity is to torture babies. If Korsgaard’s account is to have any force, then, it must have some way of introducing limits on the kind of identities that one is allowed to adopt. And indeed it does. Let us see how.

We have already seen that a law can have binding force on an agent only if the agent has some representation of the law. That is, a principle that is not endorsed by the agent has no authority over him. For instance, if I do not identify with being a good soldier, then I would feel

\textsuperscript{1} That is not to deny that they may be social constructed, but that they are adopted in a very personal way and further that they combination of these identities are amalgamation of conceptions that are unique in each individual. Thus, bits and parts of one’s identity may be constructed by society and imposed on the individual. However, one organizes these bits and pieces by its own volition, and thus makes up an identity that is completely personal and internal to the self.
no categorical urge to wax my shoes every morning. I might do it because I would get into trouble if I didn’t, but that would be only hypothetically imperative for me. For it to be categorically imperative I would have to identify myself with being a soldier.

Interestingly this is also true in case of being reflective itself. I am subject to the principles that come with being reflective, only if I identify with being reflective (Korsgaard 1996, 121). Further, if we assume as Korsgaard does, that being a human being just is being reflective, then it would only make sense to expect reflectivity of me, if I endorsed being a human. In other words, the first statement of the categorical imperative that requires of any rational being to “Act only in accordance with that maxim through which [he] can at the same time will that it becomes a universal law,” has binding force on any being insofar as it identifies itself as a rational being (Kant 1998, 4:421). To put that in language by now familiar, let us simply say, since every decision we make implies reflective thinking, it also implies that the agent identifies with being a human. Humanity, then, is a fundamental identity that precedes others (Korsgaard 1996, 121). In so far as we value our identities, if it is true that the most fundamental of our identities is our identity as a human being, then being a human is one thing we can surely claim to be universally valuable for all human beings.

Thus, there is a universal identity that defines fundamental boundaries for the kind of decisions we make, including about the other kinds of identities we adopt. For example, since it would be contra our commitment to humanity to torture little babies, then it would be immoral to torture babies, no matter what. Since you would have to commit to being a human being before committing to any other identity, you would not just do anything that that identity requires of you, because your identity as a human being, that is, a reflective agent, requires of you to value humanity over all other considerations, and that means refraining from a heinous act like torturing babies.

In the last few pages, I have examined an important consequence of the Kantian theory: Humanity is valuable unconditionally. Suddenly, it seems like Williams’ charge has regained its force. That is, if morality requires of you to value something outside of yourself unconditionally (like humanity) then isn’t that just to say that there is an impartial view from where morality is determined? Strangely the answer to that question is yes. However, what I have established in this paper is that humanity’s demands on us are not, although they certainly can seem to be, outside of ourselves. Like all our identities, humanity engages the will by being first adopted by the agent as a categorical principle. Thus, it is insofar as I identify with being a human that I am committed to not torturing little babies, for example. As such, it is true: We are bound, but we are bound as far as we choose to be. It is our humanity that requires us to value humanity, nothing else.

REFERENCES

