Consent and the Mere Means Principle

Kant’s goal in the *Groundwork to the Metaphysics of Morals* is to seek out and establish the Supreme Law of Morality. According to Kant, the Supreme Law of Morality manifests as an imperative to imperfectly rational beings like humans: the Categorical Imperative. In section two of the *Groundwork*, in an attempt to bring the Categorical Imperative closer to intuition, Kant articulates the Formula of Humanity:

\[
\text{FH} \quad \text{Act so that you use humanity, as much in your own person as in the person of any other, always at the same time as an end, never merely as a means.}
\]

The FH can be analyzed into two parts. One is an injunction to treat humanity always as an end. The other is a prohibition on using humanity as a mere means. The second is often referred to as the FH prohibition or the mere means prohibition. It has become popular to interpret this prohibition in terms of consent. The idea is that, if X uses Y’s humanity as a means and Y does not consent to it, then X uses Y’s humanity as a mere means (and therefore impermissibly according to the FH). There is then debate about the kind of consent that is relevant: possible, actual, or rational.

In this paper, I argue against this interpretation. Section one sets up the consent account. I look at the passage from the *Groundwork* in which Kant seems to appeal to lack of consent as a sufficient condition for using someone’s humanity as a mere means. I then survey some of the literature that this passage has generated and explain the differences between the different kinds of consent. Section two attacks possible and actual consent accounts on doctrinal grounds. Section three extends this doctrinal attack to rational consent accounts. Section four circles back to the original passage from the *Groundwork*. I argue that the consent account misinterprets this passage. I conclude with a quick sketch of an alternative interpretation of the FH prohibition.

Section 1. Motivating the Consent Account

After articulating the FH, Kant uses it to illustrate four thought experiments. The second of these is the most important for current purposes. Kant asks us to imagine someone who is in financial distress. The person contemplates telling a lying promise in order to get some ready money. Kant asserts that anybody in this position immediately will see that such a promise uses the promisee as a mere means. It is what comes next, Kant’s explanation of why this is so, that is crucial:

\[
\text{He whom I will to use through such a promise for my purposes cannot possibly consent to my way of acting toward him and thus himself contain the end of this action. This conflict with the principle of other humans comes clearly into view if one brings in examples of attacks on the freedom and property of others. For then it becomes clear that the transgressor of the rights of humans is of a mind to have the person of others serve merely as a means, without taking into consideration that these persons, as rational beings, should be valued always at the same time as an end, i.e. only as such beings which also must be able to contain in themselves the end of the very same action.}
\]

Kant’s reasoning in this passage suggests two related but distinct conditions that must be satisfied for an action not to run afoul of the prohibition on using others as mere means: (i) a consent condition and (ii) an end-containment condition.

There are two reasons why the consent condition might be more prominent in the current literature than the end-containment condition. First, it is unclear how to understand the end-containment condition. Consent is also a difficult concept. But consent is prevalent in everyday life and in moral philosophy. So, in the absence of an
explanation from Kant of either of these conditions, the consent condition might seem more tractable. Second, in the final sentence of the passage, Kant says that if an agent is able to contain the end of an action, then she is being valued at the same time as an end. From this it follows that, according to Kant, if an action passes the end-containment condition, it passes the FH. But there are more actions that fail to treat humanity at the same time as an end (and, therefore, fail the end-containment condition) than use humanity merely as a means (and, therefore, fail the FH prohibition). So, if we are limiting ourselves to interpreting the FH prohibition (rather than the FH as a whole), only the consent condition can be equivalent.

However, there is debate about how the consent condition should be understood. As noted in the introduction, there are three main options: possible, actual, and rational consent. O’Neill and Korsgaard advocate possible consent. According to Korsgaard,

The sense of “could not possibly” [consent]...is literal. The victims of forceful, coercive, and deceptive actions cannot consent because these actions by their nature give their victims no opportunity to consent.

There is a subtlety here that is sometimes overlooked. We can distinguish agent-centered from patient-centered accounts of consent. This distinction is frequently made in discussions of informed consent. An agent-centered account of informed consent asks whether the agent provides sufficient information to the patient. A patient-centered account asks whether the patient has sufficient information. These diverge if the patient has information that the agent did not provide or if the patient was distracted while the agent was providing information.

This distinction is also useful for possible consent. An agent-centered account of possible consent focuses on the agent’s maxim of action. A patient-centered account focuses on the condition of the patient. To see how these come apart, consider the following example from Kerstein:

Suppose that I hail a cab and ask the driver to take me to the airport. But, unknown [sic] to me, the driver cannot refuse my request; he cannot avert my use of him. For he has been hypnotized into being unshakably convinced that he cannot turn down any request I make.

Kerstein thinks that this example poses a problem for the possible consent account because (a) it is not possible for the cab driver to refuse Kerstein’s use of him, and (b) the cab driver is not being used merely as a means. But (a) is true only if we appeal to a patient-centered account. Because both O’Neill and Korsgaard appeal to agent-centered accounts, Kerstein’s example misfires.

Kleingeld advocates actual consent. According to Kleingeld,

An agent uses another person merely as a means if and only if (1) the agent uses another person as a means in the service of realizing her ends (2) without, as a matter of moral principle, making this use conditional on the other’s consent; where (3) by ‘consent’ is meant the other’s genuine actual consent to being used, in a particular manner, as a means to the agent’s end.

Cherry also advocates actual consent. The difference between actual and possible consent may be illustrated using the concept of tacit consent. An agent can give tacit consent to something when she does not protest even though she does not express consent. So, if an agent gives the appearance of tacit consent to a maxim that it is possible for her to consent to, then the action will satisfy an agent-centered possible consent account like Korsgaard’s or O’Neill’s. But, the mere appearance of tacit consent cannot suffice for an action to satisfy an actual consent account like Kleingeld’s. For example, if Kerstein acts on a maxim to seek another cab should the driver protest his request, then his maxim arguably will pass an agent-centered possible consent account. But because, unbeknownst to the
agent, the driver is not competent (because he has been hypnotized) and so the tacit consent is spurious, the action would qualify as using the driver as a mere means on Kleingeld’s actual consent account.

Papadaki, Rohlf, and Parfit advocate rational consent. Understanding rational consent requires a substantive account of rationality, and there is no standard account. Papadaki explains rational consent in terms of ends that an agent is rationally compelled to have:

An agent X does not treat Y merely as a means, if...Y’s not consenting to this sort of treatment would entail consenting to give up an end that Y is rationally compelled to have. The end in question being respecting humanity.

Rohlf argues that rational consent should be understood in terms of another of Kant’s formulations of the Categorical Imperative:

How then does Kant think we should decide what other people can rationally consent to? Kant answers this question in the formula of autonomy: on his view, we can rationally consent to all and only actions whose maxims can be willed as universal laws.

Parfit maintains that whether an agent rationally can consent to something depends on consequentialist considerations:

Whether we could rationally consent to some act depends in part on the benefits or burdens that would come to us or other people in the different outcomes that would be produced by this and the other possible acts.

Rational consent diverges from possible and actual consent insofar as it is hypothetical. I explain some of the implications of this below.

Section 2. Doctrinal Reasons for Questioning Possible and Actual Consent Accounts

In the Groundwork, Kant argues that someone who commits suicide from self-love acts contrary to the FH prohibition because she uses her own humanity as a mere means. Actions that use an agent’s own humanity as a mere means are categorized as violating owed duties to oneself. In the Metaphysics of Morals, Kant adds other offenses to this category, including lust and gluttony. That an agent can use her own humanity as a mere means is made explicit in Kant’s discussion of servility:

The human alone, regarded as person, i.e., as subject of a moral-practical reason, is raised above all price; for as such a one (homo noumenon) he is not merely to be valued as means to the ends of others, [or] indeed even his own [ends], but rather as end in itself.

This poses a problem for actual consent accounts because it signifies that an agent can consent to being used as a mere means. It poses a problem for possible consent accounts because it signifies that there are maxims to use a patient as a mere means which it is possible for the patient to consent to. Moreover, the problem is not merely theoretical. Consider what Kant says about marriage:

The man can neither desire the woman in order to enjoy her as a thing...nor can the woman give herself thereby to him without both giving up parts of their personality...i.e., without the condition of marriage.
On Kant’s account, if people who are not married have sex, they are using each other as mere means. In his discussion of lust, Kant goes further. He argues that the purpose of sex is procreation, and he asks:

During pregnancy, for example—[or] when the woman is sterile (on account of old age or illness), or if the woman finds no desire in herself for it, is it not contrary to the end of nature and thereby also duty to oneself for one or the other party, even as in unnatural lust, to make use of their sexual attributes?  

Kant suggests that there might be a permissive law allowing intercourse in such cases. But he does not take a definitive stand in the *Metaphysics of Morals*. This is part of the casuistical questions, so Kant leaves it to his readers to think through the issues for themselves. The point for present purposes is that, on Kant’s account, marriage is a necessary but (almost certainly) insufficient condition for sex not to constitute using another as a mere means. But non-procreational, extramarital sex can be consensual. So, we have a concrete counterexample to the actual and possible consent accounts of the FH prohibition: neither actual nor possible consent accounts give a sufficient condition for treating someone as a mere means.

Now some might object that Kant’s views on sex and marriage are mistaken. And they might be right. Others might reject duties to oneself wholesale. And this rejection might be justified. But both of these objections miss the point. My objection to possible and actual consent accounts is not based on the philosophical probity of Kant’s claims. Rather, my objection is based on the bare fact that Kant makes these claims, whence I conclude that the actual and possible consent accounts cannot be what Kant had in mind with the FH prohibition. There is an unstated assumption in my argument about how charitable we should be when reading Kant. I return to this assumption below (in section 4 of this paper). But, for now, the point is that there are doctrinal grounds for rejecting the possible and actual consent accounts as readings of Kant.

Some might be undeterred by this. Kant’s prohibition has taken on a life of its own. So some might reject Kant’s prohibition in favor of a Kantian prohibition. They might concede that Kant did not subscribe to an actual or possible consent interpretation. But (they might argue) actual and possible consent get the better of Kant when it comes to sex and marriage. So perhaps we should forget about Kant and see where consent leads.

However, I think this would be a mistake. Even if my foil rejects duties to oneself as incoherent, it is an empirical fact that some people are servile. And a servile person might consent to being used as a mere means to another’s ends. So I do not think that the actual or possible consent interpretations should ground a Kantian prohibition, much less Kant’s prohibition.

Alternatively, some might concede that it is possible to consent to being used as a mere means and, indeed, that servile people are disposed to do so. But (the objection continues), on Kant’s account, agents *ought* not to consent to being used as mere means. So, even if it is possible to do so and even if some agents actually do so, consenting to being used as a mere means is irrational on Kant’s account.

Now the concession that such actions are possible is fatal to both possible and actual consent accounts, and that is what I have been driving at. But this objection is nonetheless important. As noted above, rational consent accounts diverge from possible and actual consent accounts because rational consent is hypothetical. Even if an agent can and does consent to being used as a mere means, if that consent is irrational, the rational consent interpretation of the FH prohibition remains intact. Rational consent accounts face different problems.

Typically, rational consent accounts are rejected precisely because they are hypothetical and, thus, come apart from actual consent accounts. For example, in professional medical ethics, physicians are generally taught that they may not overrule a patient’s informed and considered judgment about treatment, not even if the physician thinks that this judgment is grossly irrational. However, such an objection will not work in the present context. For one thing, I am engaged in an exegetical project: my goal is to show that Kant’s mere means prohibition should not be interpreted in terms of consent, rational or otherwise, at least if we are going to remain true to Kant’s original intentions. My goal is not to show that, if the mere means prohibition is interpreted in terms of consent, then it has counterintuitive implications. For another thing, any such objection will rely on a substantive account of rationality. If this account is not Kant’s, the objection will not show even that Kant’s
prohibition, interpreted in terms of rational consent, has counterintuitive implications. This second problem recalls something I pointed out in section 1 of this paper. Different proponents of rational consent interpretations of the FH prohibition favor different accounts of rationality. So, in the next section of this paper I am going to address each account of rational consent individually before raising a general doctrinal problem for rational consent accounts. Then, in the final section, I shall return to the direct textual evidence regarding consent interpretations in order to wrap up my argument against them.

Section 3. Doctrinal Reasons for Questioning Rational Consent Accounts

Recall that Papadaki asserts that an action passes the FH prohibition if not consenting to it requires the patient to consent to give up the end of respecting humanity. In the *Metaphysics of Morals*, Kant distinguishes between duties of respect and duties of love. Duties of respect follow from the FH prohibition on using humanity as a mere means; duties of love follow from the FH enjoinder to use humanity at the same time as an end. So, rewording Papadaki’s account: an action passes the FH prohibition if not consenting to it requires the patient to consent to violate the FH prohibition. Or, perhaps more perspicuously (in contrapositive form): if an action violates the FH prohibition, then consenting to it requires the patient to consent to violate the FH prohibition.

Understood as such, Papadaki’s account has at least two problems. First, the FH prohibition appears on both sides of this conditional. So, this conditional cannot be used to elucidate the FH prohibition. Second, the addition of consent to the conditional, the only thing that stops it from being vacuously true, scuttles it. If it is not possible to consent to an action, then not consenting to it does not require the patient to do anything. It follows (*a fortiori*) that, if it is not possible to consent to an action, then not consenting to it does not require the patient to consent to give up the end of respecting humanity or to consent to violate the FH prohibition. So, if an action fails the possible consent account, it will pass this criterion. This is a problem because, although the possible consent account might not be Kant’s, and although Kant might have had good reason not to subscribe to it, it is not the case that every action that fails the possible consent account is permissible, nor is it the case that Kant would condone every such action.

Here are two things that Papadaki might say in response. First, Papadaki might point out that, when she speaks of the end of respecting humanity, she is not invoking Kant’s technical distinction between duties of respect and duties of love. Rather, what Papadaki has in mind is that an agent must consent to give up the end of respecting humanity if but only if she must consent to violate the FH. So, my paraphrase of Papadaki’s account is incorrect. It should be: an action passes the FH prohibition if not consenting to it requires the patient to violate the FH. Or, again, perhaps more perspicuously (in contrapositive form): if an action violates the FH prohibition, then consenting to it requires the patient to consent to violate the FH. Second, Papadaki might note that, if it is not possible to consent to an action, then it is (*a fortiori*) not possible rationally to consent to that action. So, the second problem I raised for Papadaki’s account is off-target.

However, I do not think that these responses work. If whether an action violates the FH prohibition is conditional on whether that action requires the patient to consent to violate the FH, then the distinction between duties of respect and duties of love is, at least *prima facie*, untenable. So, precisely because Kant thinks that this distinction is meaningful, it provides doctrinal grounds for rejecting Papadaki’s rational consent criterion. Moreover, the emendation suggested in the previous paragraph does not actually fix either of the problems I raised for Papadaki’s account. The FH prohibition is a proper part of the FH, so we are still trying to elucidate the FH prohibition with itself, only now there is the added paradox that the whole is being used to illuminate the part. And while I concede that, intuitively, any action that fails the possible consent criterion (*a fortiori*) fails the rational consent criterion, I do not see how this obviates the second problem I raised for Papadaki’s account, which is based on a close-reading of her text—if anything, this only suggests that Papadaki’s account cannot be used to give an adequate understanding of rational consent.

As noted above, Rohlf maintains that “we can rationally consent to all and only actions whose maxims can be willed as universal laws.” But this faces a problem analogous to the one raised against Papadaki about the
distinction between duties of respect and duties of love. Assuming that the different formulations of the Categorical Imperative are equivalent, Rohlf’s account commits him to the following three biconditionals: (1) an action violates the FH prohibition if and only if it violates the rational consent condition; (2) an action violates the rational consent condition if and only if it violates the condition that maxims can be willed as universal laws; and (3) an action violates the FH if and only if it violates the condition that maxims can be willed as universal laws. But these three biconditionals entail a contradiction when conjoined with (4) there are actions that do not violate the FH prohibition that violate the FH. And as argued in section 1 of this paper, Kant is committed to (4). So, at least one of the biconditionals in Rohlf’s account must be rejected.\footnote{37}

Parfit’s understanding of rationality entails that whether an action passes the rational consent condition depends on the action’s consequences: the benefits and burdens that would accrue to us and to others given the alternative actions available. This is not the place to engage in a full-scale explanation of why Kant’s ethics is inconsistent with consequentialism.\footnote{38} I note merely that Kant never makes an argument of the kind Parfit is attributing to him, neither in the Groundwork lying promise example, nor anywhere else in the Groundwork or his other ethical writing, from the Critique of Practical Reason to the Metaphysics of Morals. Indeed, in his discussion of the duty of benevolence in the Groundwork, Kant suggests that the world would be a better place if everyone were an egoist (provided that the current tendency to cheat were not so prevalent)—but he then says there is a duty to be benevolent all the same.\footnote{39} That is, the only consequentialist reasoning in this section of the text actually militates against a duty of benevolence rather than in favor of one.\footnote{40} So, Parfit’s theory of rational consent has scant textual support and, indeed, seems not to be in the spirit of Kant’s ethics.\footnote{41}

I want to close out this section with one last point, a general problem for rational consent interpretations of the FH prohibition. Kant famously subscribes to the idea that morality is a form of rationality. It follows that the FH, as a formulation of the Categorical Imperative, is intended to be a standard of rational action. At first blush, this might seem like a point in favor of the rational consent interpretation. But now consider: (1) an action is rational only if it is in accordance with the FH prohibition; and (2) according to the rational consent interpretation, an action is in accordance with the FH prohibition if and only if its patient rationally can consent to it. Telescoping these, we get: an action is rational only if its patient rationally can consent to it. There are two (at least \textit{prima facie}) problems with this, the same two problems that we saw above for Papadaki’s account. One is that rationality appears on both sides of the conditional. To draw this out, note that consent is an action, so a violation of regularity looms. The other problem is that, if consent has a substantive role to play here, it will undermine the interpretation. To see why, note that, if there are actions that a patient is unable rationally to consent to but that rationally can be performed with/on/to her, then the biconditional in (2) is false. Let me try to make this second problem clearer. Proponents of the rational consent interpretation take rationality to play a substantive role in their interpretation of the FH prohibition. That is, as we saw at the end of section 2 of this paper, it is possible for people (e.g., servile people) to consent to be used as a mere means even though they ought not to do so. But, the argument goes, people cannot \textit{rationally} consent to such actions. So, the mere fact that people can consent to be used as a mere means is not a problem for the rational consent interpretation. That is what I mean when I say that rationality plays a substantive role in the interpretation. And what I am pointing to now is the correlate of this: if consent, like rationality, is playing a substantive role, then there are actions that a patient is unable to consent to and, a \textit{fortiori}, unable rationally to consent to even though they rationally can be performed with/on/to her—whence it follows that, if consent is playing a substantive role in the rational consent interpretation, then the biconditional in (2) is false. Thus, the fact that Kant takes morality to be a form of rationality, so far from being a point in favor of the rational consent interpretation, actually militates against it.

\textbf{Section 4. What We Should Have Said Is...Nothing}

As noted in section 2 of this paper, my doctrinal arguments against consent accounts rely on a principle of charity. I argued that these interpretations of the FH prohibition create problems for what Kant says elsewhere. But Kant is human: he sometimes contradicts himself, changes his mind, or fails to understand his own ideas. So maybe
something like this happened with Kant’s mere means prohibition. That is, perhaps Kant did intend for this prohibition to be interpreted in terms of consent, and perhaps anything he said that suggests otherwise results merely from a misunderstanding or mistake on his part. If so, then my attacks in the previous sections do not connect. In order to lay this objection to rest, I propose to return to the lying promise example that inspired the consent condition in the first place.

Recall the following quotation from section 1 of this paper: ”[H]e whom I will to use through such a promise for my purposes cannot possibly consent to my way of acting toward him and thus himself contain the end of this action.” The word that I have translated as ‘consent’ is *einstimmen*. *Einstimmen* is usually translated here as ‘agree.’ In English, this connotes mutual consent. But a more accurate translation would be something like ‘align’ or ‘attune’, which do not have such connotations. *Einstimmen* does not mean consent, nor does it mean anything that entails, presupposes, or is analyzable in terms of consent. As evidence for this, consider the following 20 passages in which Kant uses the term as verb, adjective, and noun:

1. For if they at some time...wake up completely, i.e. open their eyes to a view which does not close out *Einstimmung* with the understanding of other humans, then none of them will see anything which should not appear to everyone else at the same time, with the light of their proofs, as apparent and certain.  
2. [E]ven if [universal assent] were to happen to hold, this nevertheless could not yet give a proof of the concordance [*Übereinstimmung*] with the object; rather objective validity alone accounts for the ground of a necessary universal *Einstimmung*.  
3. Even though otherwise a universal law makes everything *einstimmig*, here, if one willed to give the maxim [to make happiness the determining ground of one’s will] the universality of a law, exactly the furthest opposite of *Einstimmung*...would follow.  
4. [I]f everyone were to allow himself to cheat when he believed it to be to his advantage, or held himself authorized to foreshorten his life so soon as a complete weariness of it befell him, or looked upon others’ necessity with complete indifference, and you belonged to such an order of things, would you be therein with the *Einstimmung* of your will?  
5. [W]ithout *Einstimmung* with the moral law, [an agent’s opinion of his personal worth] will be reduced to nothing.  
6. [E]very will...is limited to the condition of *Einstimmung* with the autonomy of the rational being...thus to use this being never merely as a means but rather at the same time as an end.  
7. The judgment of taste does not itself postulate everyone’s *Einstimmung*...it only expects this *Einstimmung* of everyone.  
8. [T]he object could give it [the imagination] just such a form, which contains a composition of the manifold, as the power of imagination, if it were left free to itself, would produce in *Einstimmung* with the lawfulness of the understanding in general.  
9. [T]he furtherance of happiness in *Einstimmung* with morality [is the final end of practical reason].  
10. [T]here are] principles of the great societies, called states...which no philosopher has yet brought into *Einstimmung* with morality.  
11. Appropriation (*approbatio*) as act of an external universal lawgiving will (in the idea), through which everyone is bound to *Einstimmung* with my will.  
12. [A right against a person must conform to] the principle of *Einstimmung* of the freedom of my will with the freedom of everyone.  
13. Even if a civil society, with all members’ *Einstimmung*, dissolves itself...the last murderer found in prison must be executed first.  
14. But the late Michaelis...only gave his rational propositions, through true or imagined *Einstimmung* with others’ (perhaps poets’ and orators’) judgments, clarification and confirmation.  
15. For this is the touchstone of the rightfulness of any public law. If namely this [law] is so conditioned that it would be impossible for an entire people to give its *Einstimmung* to it...then it is not rightful.
Differences of languages and religions...with increasing culture and the gradual approach of humans to greater Einstimmung in principles, lead to concordance in a peace which is brought forward and secured...through equilibrium in liveliest competition.59

Only in relations of qualities, as my understanding should form entirely a priori concepts of things, with which the things should necessarily einstimmen, how it should produce real principles about their possibility with which experience must truly einstimmen and which nevertheless are independent of it—this question always leaves behind a darkness in relation to our capacity of understanding from whence this Einstimmung with the things themselves comes to it.60

The end goal of aesthetic cognition is truth and goodness. Hence they einstimmen to understanding and will through subjective means.61

The existence of a merely ethical being without happiness has indeed, for an observer, the greatest worth, but not for the subject itself...For I can as well say: I must be truthful, even should fortune deny me any favor; this however is valid only insofar as I am there and live, exist as a good being. But I do not know why I should there there merely to act. Even the very will (in the idea) which has commanded this to me, why has he required my existence. I cannot einstimmen thereto.62

The end goal of aesthetic cognition is truth and goodness. Hence they einstimmen to understanding and will through subjective means.

Some of these passages work with einstimmen taken to mean consent. For example, passage 4 could be read as asking whether you would consent to be part of a world in which anybody who believed it to be to her advantage would cheat, etc. Similarly, passage 15 could be read as saying that, if it would be impossible for an entire people to consent to a law, then that law is not rightful.

But most of these passages (1-3, 5-11, 14, and 16-20) would be, at best, strained and, in many cases, positively nonsensical with einstimmen taken to mean consent.64 If we look at passage 1, then we see that it makes no sense to say that a view closes out "consent with the understanding of other humans." That is, the faculty of understanding does not engage in consent, and the consent of other humans is not what is wanting here. Kant is talking about the alignment of understanding, an alignment that could be referred to as agreement if we are speaking figuratively—but not consent. Skipping to passage 5, we see Kant saying that an agent’s estimation of her own worth will be reduced to nothing "without Einstimmung with the moral law"—which arguably gives a category mistake if we try to plug in ‘consent’. Similarly, in passage 6, when Kant talks of a gradual approach of humans "to greater Einstimmung in principles," this seems, at least to me, to be incoherent if we interpret it as ‘consent’—and so on down the line: in passage 18, can they consent to understanding and will through subjective means? To my mind, this is strained at best, and the same can be said of passage 20, when nature is supposed to be einstimmend with freedom.63

This undercuts the original motivation for the consent condition. The argument I am making is not that interpreting Einstimmen (and its derivatives) as consent always results in incoherence and, therefore, Einstimmen never should be interpreted in this way. Rather, the argument I am making is that interpreting Einstimmen as consent often results in incoherence and, therefore, there should be a presumption against doing so. There is then a further exegetical complication: Kant does not use a consent condition to derive duties in the Metaphysics of Morals even though this work contains his most extensive and detailed discussion of our general duties. These two things (i.e., the positive textual evidence against interpreting Einstimmen in terms of consent, and the lack of evidence of Kant deploying a consent condition elsewhere) make a strong case, I think, for rejecting a consent interpretation of the FH prohibition. So, what should be put in its place?

There are some alternatives on offer. Recent work by Seymour Fahmy and Sticker is instructive.65 However, I want to provide a brief sketch of a positive proposal of my own, with the caveat that I cannot, and do not intend to, offer anything like a full defense or exposition of it here.

Derivations of duties using the consent condition have the following structure:

1. The FH prohibits using agents as mere means.
2. X’s action uses an agent as a mere means if but only if X’s action fails the consent condition.
3. X’s action fails the consent condition.
4. Therefore, X’s action is prohibited by the FH prohibition.

So the question I am asking is: what should be put in the place of 2? And the answer I advocate is: nothing. That is, I think this four step derivation should be replaced with a three step one:

1. The FH prohibits using agents as mere means.
2. X’s action uses an agent as a mere means.
3. Therefore, X’s action is prohibited by the FH prohibition.

I do not think that Kant took there to be necessary or sufficient conditions to elucidate 2, and I think he was right about this although, as noted above, I cannot argue for that here. But let me say a little bit more to flesh this out.

I think that whether an agent’s action uses someone as a mere means is a hermeneutic question about the meaning of the action. Answering it requires a thorough understanding of the agent’s action, which in turn requires an understanding of the agent’s culture, background beliefs, character, circumstances, and intentions. This can be illustrated by appeal to practices like dueling, child labor, capital punishment, or the treatment of the elderly. Cultural norms, which inform agents’ background beliefs and intentions, have differed wildly regarding these in different places and at different times, and that, in turn, can inform whether a specific instance of these practices is in conformity with or in violation of the FH prohibition. An analogy here might be instructive. The Supreme Court of the United States now interprets the eighth amendment, prohibiting cruel and unusual punishment, in terms of the evolving standards of decency. Thus, for example, when capital punishment was restricted to adults in 2005, it was so on the grounds that, if defendant was a juvenile at time of offense, subjecting her to the death penalty would be cruel and unusual under the evolving standards of decency test. In the same way, I am arguing that whether an action constitutes using someone as a mere means can be informed, at least in part, by background cultural norms.

I do think that consent can play a relevant role in explaining why a given action uses another as a mere means. But the point I want to make is that consent can play a role only in the explanation: it is not an equivalent, necessary, or sufficient condition on Kant’s account, and that is why I propose to eliminate, rather than replace, premise 2.

Some might object that I am advocating moral particularism. Kant’s ethics is based on principles. Principles have universal ambit. But (the objection goes) I am proposing to leave agents to some sort of rational intuition about particular actions; the only principle that I allow is that there is no principled way of determining whether an action is in conformity with the FH prohibition. Not only is this not Kant, it is probably not coherent.

But this objection gets everything wrong: it gets Kant wrong, it gets the consent account wrong, and it gets me wrong.

Kant maintains that the faculty of judgment is required in order to apply any principle or concept to experience. The faculty of judgment can make mistakes, and it can be trained (well or poorly). Some people have good judgment, meaning that they readily and generally correctly subsume concrete particulars under universal principles or concepts. Others have bad judgment, meaning that they do not enjoy this facility. The principle in this case is: <using someone’s humanity merely as a means>. And the point is that, on Kant’s account, subsuming any particular action under this principle is going to be a matter of judgment. So, this objection about particularism gets Kant wrong.

Now suppose for the sake of argument that some version of the consent account is correct. The consent account does not get around the need for judgment. I have tried to make that clear by contrasting the 4-step consent derivation of duties with the 3-step non-consent derivation of duties: as far as the objection now under consideration goes, there is no material difference between these derivations. The consent account introduces another principle between <using someone’s humanity merely as a means> and the particular actions that are subsumed under this principle. But there is no algorithmic, failsafe method for applying <failing the consent
account> to particulars any more than there is for applying <using someone’s humanity merely as a means>. So, this objection about particularism gets the consent account wrong.

Finally, the account that I am advocating is not a version of moral particularism, nor does it run contrary to Kant’s principlism. The account that I am advocating is built on a principle that is meant to be a version of the Supreme Law of Morality as applied to humans: the FH. This is the opposite of moral particularism, and it is precisely the principle that Kant uses most frequently in his derivations of duties in the *Metaphysics of Morals*. Which takes me to the last thing I would like to say in favor of my interpretation of the FH prohibition.

If there were any condition like the consent condition, it would be easy to find in the text, and it would be easy to determine which version of it is correct. In the second half of the *Metaphysics of Morals* Kant derives a series of owed duties, usually using the FH prohibition. But here is the thing: these derivations proceed directly from the FH prohibition, not from an intermediate principle having to do with consent or, for that matter, anything else. More, these discussions are usually capped with casuistical questions, in which Kant proposes various possible exceptions to the general duties he has just derived. What that indicates is that the general duties in the *Metaphysics of Morals* are not exceptionless; Kant is not trying to articulate there, any more than he was in the lying promise example from the *Groundwork*, necessary or sufficient conditions for violating the FH or the mere means prohibition. And even if we disagree with him about some of the duties he came up with (like the duties he articulates regarding sex and marriage), perhaps his overall strategy is the way to go. If, as I am arguing, this strategy is the one we should follow, not only textually but philosophically, we still can have absolute and universal prohibitions. In particular, there is still an absolute and universal prohibition on using others as mere means. The point is only that determining what other absolute and universal prohibitions there are must proceed directly through the FH and the mere means prohibition rather than via an intermediate principle about consent or some other concept.

**Conclusion**

In this paper, I have attacked consent accounts of Kant’s mere means prohibition. In section 1, I examined the text from the *Groundwork* that is used to motivate these accounts, and I briefly set out the different kinds of consent that may be found in the literature: possible, actual, and rational consent. In section 2, I advanced doctrinal grounds for doubting that possible or actual consent can provide an interpretation of the mere means prohibition that is true to Kant’s original intentions, and in section 3 I extended this attack to rational consent. In section 4, I returned to the text from section 1 and argued that the original motivation for consent accounts is based on a mistranslation. I then briefly sketched a proposal for what I would put in place of the consent accounts: nothing.
Bibliography


Pascoe, Jordan. 2012. ”Kant and Kinky Sex.” In What Philosophy Can Teach You About Your Lover, ed. Sharon M. Kaye: 25-36. Open Court


Patrone, Tatiana. 2018. ”Treating Others as Means, but Not Merely as Means.” Ethical Perspectives, Vol. 25, Number 1: 61-86.


1 GMS, AA 04: 392.03-04. All citations are to the standard Prussian Academy editions of the works of Immanuel Kant. All translations, except where otherwise noted, are my own, although I have consulted, where possible, the Guyer/Wood Cambridge Editions of the Works of Immanuel Kant.

2 GMS, AA 04: 412.26-413.11. The implications of this are explored in (Kahn, 2021).

3 GMS, AA 04: 429.10-12, emphasis omitted.

4 Calling it "the" FH prohibition is somewhat misleading. To say that you ought to treat humanity, whether in yourself or any other person, always as an end is also, arguably, a prohibition. That is, the FH does not command us to perform every action that treats humanity in every person as an end. Such a command would be absurd. Rather, the FH forbids us to perform any action that fails to treat humanity in any person at the same time as an end. So, the mere means prohibition is (merely) part of a broader prohibition, and that broader prohibition perhaps more rightly could be called "the" FH prohibition. I would like to thank an anonymous reviewer from The Journal of Value Inquiry for pressing me to clarify this.

5 GMS, AA 04: 429.33-430.09.

6 See Formosa: "According to the above passage, I use another as a mere means if, first, the other cannot possibly agree or consent to my way of acting toward her or if, second, the other cannot possibly contain in herself the end of the very same action" (2014, 57).

7 According to Parfit, "We treat people as ends, Kant claims, and not merely as a means, if we deliberately treat these people only in ways to which they could rationally consent" (Parfit, 2011, vol. I, p. 218). As may be inferred from this quotation, Parfit interprets the consent condition in terms of rational consent. So Parfit’s assertion may be paraphrased: if an action passes the consent condition, then it treats people as ends. But Parfit is mistaken: Kant does not claim this. Kant claims only that if an action passes the end-containment condition, then it treats people as ends, not that if an action passes the consent condition, it treats people as ends. As such, Parfit’s assertion could be attributed to Kant only if Kant claimed that passing the consent condition entails passing the end-containment condition. But Kant does not claim this either. What Kant does claim, in the first sentence of the passage above, is the converse of this: failing the consent condition entails failing the end-containment condition—exactly what we would expect if the consent condition is equivalent to the mere means prohibition whereas the end-containment condition is equivalent to the FH.

8 Patrone suggests that, in the FH, "’merely’ modifies ‘using’ rather than ‘means’" (Patrone, 2018, 75). On the basis of this she ascribes the following two claims to Kant: (1) if you are using someone as a means, you are using her merely as a means, and (2) if you are not using someone merely a means, your actions are permissible (Patrone, 2018). If Patrone is right about this, then the sentence to which this footnote is appended is false. However, neither (1) nor (2) should be ascribed to Kant, and Patrone’s reading of the FH is mistaken.

9 Regarding (i): this would make it difficult to understand the phrase ‘always at the same time’ in the FH. Indeed, Kant is explicit in the Metaphysics of Morals that there are conditions in which, because I am treating a person at the same time as an end, I "can make use of a person exactly as of a thing, nevertheless without abridgment of their personality” (MS, AA 06: 659.20-22; I owe this reference to Kleingeld, 2020a, 403). Regarding (a): in the Groundwork Kant asserts that

...in relation to contingent (meritorious) duty to oneself it is not enough that the action not conflict with the humanity in our person, as end in itself, it must also harmonize therewith. (GMS, AA 04: 430.10-13)

So Kant thinks an agent can fail to fulfill his duties to himself even if he does not use his own humanity as a mere means. And Kant makes a similar claim about duties to others a few lines later (GMS, AA 04: 430.19-24). This is repeated in the Metaphysics of Morals when Kant argues that, because I "nevertheless can be indifferent” to the needs and wants of others, I can fail in my duty to treat persons as ends even if I use neither myself nor anyone else as a mere means (MS, AA 06: 395.15-21, esp. 19). Regarding Patrone’s reading of the FH: Patrone’s suggestion that ‘merely’ modifies ‘using’ rather than ‘means’ overlooks the fact that there is only one verb that governs both the mere means clause and the end clause of the FH.


10 (Korsgaard, 2004, 80n10).

11 (Richards, 2009, pp. 296-297).

12 (Kerstein, 2013, pp. 73-74). The example also appears in his (2009, p. 173).

13 A few brief comments: I am not sure that hypnotism can work in this way. Moreover, being convinced that one cannot turn down any request is not the same as not being able to turn down any request. Finally, as the example is constructed, it is impossible for the taxi driver not to consent rather than impossible for the driver to consent. But these details can be overlooked for present purposes.
There is an important class of cases of violence that seem to involve the use of others merely as a means, although the nature of the action does not make it impossible, for the person who is used, to consent. Consider, for example, the action of using a healthy person’s organs, thereby killing him, in order to save the lives of five others. Does this constitute a case of using the person merely as a means? On Korsgaard’s account, the question here is not whether the person does or would consent, but ‘strictly’ whether the person can consent given the nature of the action. In contrast to the false promising example, here the nature of the action does not make it impossible for the victim to consent. If, in a spirit of radical altruism, he did consent, the agent’s action would still be properly described as using his organs to save five others, thereby killing him. It follows from Korsgaard’s account of the prohibition, therefore, that organ harvesting does not qualify as using the victim “merely as a means”, not even when the victim does not or would not consent. (Kleingeld, 2020a, p. 395)

However, this does not work. On Korsgaard’s account, the nature of an action is determined by the agent’s maxim. So organ harvesting does qualify as using the victim merely as a means on Korsgaard’s account if the victim has been coerced or forced. An organ harvester acting in accordance with a maxim that passes the possible consent account is not performing the same kind of action as an organ harvester acting in accordance with a maxim that does not pass the possible consent account (and an organ harvester who proceeds even in light of the victim’s refusal is almost certainly acting on a maxim that does not pass the possible consent account).

It is worth noting that, although Kleingeld says she adopts an “agent-focused actual consent interpretation,” the way that she spells out her account makes it clear that it is patient-centered, at least as I am using the terms (Kleingeld, 2020a, p. 404, emphases omitted). For example, consider the following:

...in order to be able to give genuine consent, the person whom you want to use as means needs to be mentally competent and mature enough, and the person also needs to have and understand the relevant information. The person needs to know which end you want him to serve, how you plan to use him, what this will require of him and so on. (Kleingeld, 2020a, p. 404)

That is, Kleingeld says in this passage that the person needs to how you plan to use her, not that you need to tell her how you plan to use her. If, unbeknownst to you, the agent is not paying attention when you tell her what you are going to do, then an agent-centered account will come apart from a patient-centered one like Kleingeld’s.

There is a subtlety about Parfit that is worth noting here. Parfit advances his rational consent test as an interpretation of Kant’s remarks about the lying promise example. But Parfit argues that Kant’s mere means prohibition should be understood differently, as prohibiting action that involves using someone as a means while simultaneously regarding her as a tool, where an agent regards someone as a tool if but only if he does not pay any heed to the person’s interests. Parfit then argues that Kant’s mere means prohibition, unlike the rational consent condition, has no explanatory value because (i) if I pay only the slightest heed to a person’s interests, this will not change whether my action is wrong even though it will change whether the action passes the prohibition, and (ii) if my action is permissible (as when ordering coffee from a barista), then how I regard a person is irrelevant (Parfit, 2011, vol. 1, esp. 231-232).

According to Cherry, “we do not treat someone merely as a means if he consents to be so treated” (Cherry, 2005, p. 98). I owe this reference to Kerstein (2013, 85n1).

Kleingeld tries to distinguish between possible and actual consent in a different way:

...in order to be able to give genuine consent, the person whom you want to use as means needs to be mentally competent and mature enough, and the person also needs to have and understand the relevant information. The person needs to know which end you want him to serve, how you plan to use him, what this will require of him and so on. (Kleingeld, 2020a, p. 404, emphases omitted). For example, consider the following:

...in order to be able to give genuine consent, the person whom you want to use as means needs to be mentally competent and mature enough, and the person also needs to have and understand the relevant information. The person needs to know which end you want him to serve, how you plan to use him, what this will require of him and so on. (Cherry, 2005, p. 98).

However, this does not work. On Korsgaard’s account, the nature of an action is determined by the agent’s maxim. So organ harvesting does qualify as using the victim merely as a means on Korsgaard’s account if the victim has been coerced or forced. An organ harvester acting in accordance with a maxim that passes the possible consent account is not performing the same kind of action as an organ harvester acting in accordance with a maxim that does not pass the possible consent account (and an organ harvester who proceeds even in light of the victim’s refusal is almost certainly acting on a maxim that does not pass the possible consent account).

It is worth noting that, although Kleingeld says she adopts an “agent-focused actual consent interpretation,” the way that she spells out her account makes it clear that it is patient-centered, at least as I am using the terms (Kleingeld, 2020a, p. 404, emphases omitted). For example, consider the following:

...in order to be able to give genuine consent, the person whom you want to use as means needs to be mentally competent and mature enough, and the person also needs to have and understand the relevant information. The person needs to know which end you want him to serve, how you plan to use him, what this will require of him and so on. (Kleingeld, 2020a, p. 404)

That is, Kleingeld says in this passage that the person needs to how you plan to use her, not that you need to tell her how you plan to use her. If, unbeknownst to you, the agent is not paying attention when you tell her what you are going to do, then an agent-centered account will come apart from a patient-centered one like Kleingeld’s.

Parfit advances his rational consent test as an interpretation of Kant’s remarks about the lying promise example. But Parfit argues that Kant’s mere means prohibition should be understood differently, as prohibiting action that involves using someone as a means while simultaneously regarding her as a tool, where an agent regards someone as a tool if but only if he does not pay any heed to the person’s interests. Parfit then argues that Kant’s mere means prohibition, unlike the rational consent condition, has no explanatory value because (i) if I pay only the slightest heed to a person’s interests, this will not change whether my action is wrong even though it will change whether the action passes the prohibition, and (ii) if my action is permissible (as when ordering coffee from a barista), then how I regard a person is irrelevant (Parfit, 2011, vol. 1, esp. 231-232).

I do not accept Parfit’s interpretation of Kant’s mere means prohibition. But bracketing that, there are at least three problems with Parfit’s argument. First, a moral principle need not explain why all (or even many) wrong actions are so for it to have explanatory value. Second, if how I regard a person is part of my maxim of action, then it is relevant in determining whether my maxim and, thus, my action is permissible. Third, Parfit’s attempt to divorce the lying promise example remarks from the FH prohibition disagrees with the text.
According to Pascoe, (1) Kant “rejected the idea that sex was about procreation,” and (2) Kant thought that, even in marriage, “sex is really never morally permissible” (Pascoe, 2012, p. 32). I think both of these attributions are mistaken. Against (1): in the *Metaphysics of Morals* Kant states that “nature’s purpose, in the cohabitation of the sexes, is procreation” (MS, AA 06: 426.02-03; see also 424.29-424.05). Against (2): also in the *Metaphysics of Morals*, Kant notes that “bodily union of both sexes” can be “permissible” in marriage (MS, AA 06: 425.16-17). More, the casuistical questions referred to in the paragraph to which this note is appended would make little sense if Kant thought that all sex (or even all sex that takes place “without consideration of this [procreative end of nature]” (MS, AA 06: 426.05, emphasis omitted)) is impermissible.

See (Bruers, 2016).

Alternatively, some might object that Kant had possible or actual consent in mind when presenting the mere means prohibition but that he misapplied that formula when he talks about sex and marriage. This objection takes us into something I gesture to one paragraph up: assumptions regarding how charitable we should be when reading Kant. As such, I address it more completely in section 4 of this paper. I would like to thank an anonymous reviewer from *The Journal of Value Inquiry* for pressing me to clarify this.

O’Neill attacks actual consent on similar grounds. She points out that, if Marx is correct, then actual consent can be the result of false consciousness and, when it is so, it is spurious (O’Neill, 1995, pp. 107-108). O’Neill does not notice that this also poses a problem for possible consent.

(Kleingeld, 2020a, p. 404).

A qualifier is needed here: in contemporary professional medical ethics this is so. However, at least in the Western world, this does not always seem to have been the case.

I would like to thank an anonymous reviewer from *The Journal of Value Inquiry* for suggesting this.

This last point can be reformulated as a dilemma using a lying promise example. Suppose that Abba tells Baab a lying promise in order to get some ready money. But suppose that, unbeknownst to Abba, Baab knows what is going on and wants to help Abba. If Baab gives Abba the money, is the possible consent account vitiated because Baab has consented to being used as a mere means? Proponents of possible consent point out that, although Baab has consented to give Abba the money, Baab has not consented to Abba’s maxim (which is to deceive Baab) and, indeed, it is impossible for Baab to do so. If this is correct, then it is vacuously true that, if Baab consents to Abba’s maxim, Baab is not required to give up the end of respecting humanity, whence it follows that Papadaki’s rational consent criterion is unable to account for the very example in Kant’s text that inspired it. Alternatively, if this is incorrect and Baab is able to to consent to Abba’s maxim, then although it is certainly possible to imagine scenarios in which such consent would require Baab to give up the end of respecting humanity, it also seems possible to imagine scenarios that would not require this—in which case, once again, Papadaki’s rational consent criterion is unable to account for the very example in Kant’s text that inspired it.

(Rohlff, 2009, p. 514).

An additional issue: as noted in the introduction of this paper, Kant introduces the FH in order to bring the Categorical Imperative closer to intuition. So, if we need help understanding the FH prohibition, seeking that help from other formulations of the Categorical Imperative does not seem like Kant’s strategy. Rather, the other formulations should be illuminated by means of the FH. This is borne out by the fact that most of the duties discussed in the *Metaphysics of Morals* appeal to the FH.

But see (Timmermann, 2005).

CMS, AA 04: 423.17-35.

Although it is perhaps worth noting that Kant does not suggest that a world full of egoists would be the best of all possible worlds (and, indeed, such a conclusion would be contradicted by his later claims about the highest good); rather, he is arguing that a world full of (non-cheating) egoists would be better than the current one.

Parfit might concede this without objection: his primary goal is not exegetical. See note 38 above.

CMS, AA 04: 429.03-430.01.
There is no entry for *einstimmen* in Caygill’s (2000) or Schmid’s (1795) Kant dictionaries.

43 TG, AA 02: 342.17-21.
44 KpV, AA 05: 013.02-05.
45 KpV, AA 05: 028.07-12.
46 KpV, AA 05: 069.25-30.
47 KpV, AA 05: 078.33-34.
48 KpV, AA 05: 087.21-27.
50 KU, AA 05: 240.31-241.03.
51 KU, AA 06: 451.15-16.
52 RGV, AA 06: 034.09-11.
53 MS, AA 06: 259.01-04.
54 MS, AA 06: 271.11-14.
55 MS, AA 06: 333.18-21.
56 SF, AA 07: 008.24-30.
57 TP, AA 08: 297.20-24.
58 ZeF, AA 08: 367.22-29.
61 Refl, AA 18: 547.25-548.06.
62 VAMS, AA 23: 375.27-29.

63 The way Kant’s correspondents use the term suggests that Kant’s use was not idiosyncratic (e.g., Br, AA 11: 243.17-22).

64 (Seymour Fahmy, 2021, esp. pp. 5-8) and (Sticker, 2021, esp. section 2). Seymour Fahmy follows Hill, who sets out his account in a series of publications including his (1980; 2003; and 2012, chapter 8).

65 I am indebted in the following to (Wood, 1999, chapter 5 section 8).

66 I would like to thank an anonymous reviewer from *The Journal of Value Inquiry* for pressing me to clarify this.