

# Defending the possible consent interpretation from actual attacks

Samuel Kahn

Indiana University-Purdue University, Indianapolis

---

**Abstract:** In this paper, I defend the possible consent interpretation of Kant's formula of humanity from objections according to which it has counterintuitive implications. I do this in two ways. First, I argue that to a great extent, the supposed counterintuitive implications rest on a misunderstanding of the possible consent interpretation. Second, I argue that to the extent that these supposed counterintuitive implications do not rest on a misunderstanding of the possible consent interpretation, they are not counterintuitive at all.

**Keywords:** Kant, Kant's practical philosophy, Kantian practical philosophy, Kant's ethics, Kantian ethics, Formula of Humanity, possible consent, possible consent interpretation

In Pallikkathayil's "Deriving Morality from Politics: Rethinking the Formula of Humanity," she has a lot of interesting and insightful things to say about Kant's practical philosophy. Her paper has two main goals: (1) to discredit the so-called possible consent interpretation of Kant's Formula of Humanity (FH) and (2) to advance her own interpretation of FH.

To accomplish her first goal, Pallikkathayil points out that according to the possible consent interpretation, FH rules out actions which depend on force, coercion or deception for their nature.<sup>1</sup> She argues that this creates exegetical and philosophical problems. Exegetically, Kant thinks that sometimes such actions are permissible: Pallikkathayil points out that at various places in the *Doctrine of Right* Kant condones such actions.<sup>2</sup> Philosophically, Pallikkathayil

<sup>1</sup> Japa Pallikkathayil, "Deriving Morality from Politics: Rethinking the Formula of Humanity," *Ethics* 121.1 (2010), pp. 116-147, p. 118.

<sup>2</sup> *Ibid.*, pp. 119-124.

argues that any ethical theory that rules out all such actions would be making a big mistake.<sup>3</sup>

To accomplish her second goal, Pallikkathayil gives an overview of Kant's political philosophy and, in particular, of Kant's views about our rights and why it is necessary to exit the state of nature and enter into the civil condition.<sup>4</sup> Pallikkathayil points out that our rights are indeterminate and unenforceable in the state of nature; they become determinate and enforceable only in the civil condition.<sup>5</sup> Further, she argues that an agent (X) treats another agent (Y) as a mere means if but only if either (1) X violates Y's rights or (2) X expresses the denial of the claim that Y has equal practical standing in virtue of Y's humanity.<sup>6</sup> Because rights become determinate only in the state and because violating someone's rights is treating that person as a mere means, it follows immediately that one can determine what counts as treating someone as a mere means only in the civil condition and only by appeal to the legal system in place.

Now I would like to respond to the first part of Pallikkathayil's project. There are two main proponents of the possible consent interpretation of FH: O'Neill and Korsgaard. In what follows, I shall argue for the following two theses: (1) Pallikkathayil's attack on the possible consent interpretation does not work against O'Neill, for O'Neill's version of this interpretation is flexible enough to allow some kinds of deception, force and coercion and (2) Pallikkathayil's attack on the possible consent interpretation does not work against Korsgaard, for the attack simply does not address Korsgaard's response to the idea that it might be counterintuitive to say that all actions that rely on force, coercion and deception for their nature are impermissible.

### *1. O'Neill and possible consent*

As noted above, Pallikkathayil's attack on the possible consent interpretation begins with the claim that it rules out all actions that rely on force, coercion or deception for their nature. The idea is that force, coercion and deception in principle preclude consent. At an intuitive level, we can see this by means of an example. Suppose you plan to tell me a lying promise in order to get 20 dollars but I know better. I can consent to give you the money (perhaps without even letting you know that I am onto you). But I cannot consent to the deceit: if we try to arrange it so, it ceases to be deceit.<sup>7</sup>

<sup>3</sup> *Ibid.*, pp. 124-125.

<sup>4</sup> *Ibid.*, section II.

<sup>5</sup> *Ibid.*, section II.D.

<sup>6</sup> *Ibid.*, p. 141.

<sup>7</sup> See, e.g., Onora O'Neill, "Between Consenting Adults," *Philosophy and Public Affairs* 14.3 (1985), pp. 252-277, p. 262: "It does not follow from this that nothing done in acting on a

Similar things, it is argued, can be said about force and coercion. Lurking in the background of such examples is a conceptual analysis of deceit, force and coercion that shows that to deceive, force or coerce someone into something just is to moot the issue of consent. This conceptual analysis is not provided, and that might be because it is difficult to pin down exactly why and how consent is precluded. Regardless, O'Neill seems to concede the starting point of Pallikkathayil's attack when she begins talking about the notion of possible consent:

It is plausible to think that when we act in ways that would always preclude genuine consent or dissent, we will have used others. For example, if we coerce or deceive others, their dissent, and so their genuine consent, is in principle ruled out.<sup>8</sup>

However, O'Neill quickly backs down from any universal claims about possible consent and deception, coercion and force. She argues, on the one hand, that on her interpretation, the Formula of Humanity applies only to agents' fundamental maxims; it does not apply to every intentional aspect of an action.<sup>9</sup> For my purposes here, what this means is that the possible consent interpretation does not rule out specific actions that seem to preclude consent unless those actions are based on fundamental maxims that themselves rule out the possibility of consent. O'Neill also argues, on the other hand, that understanding what makes genuine consent possible is a difficult thing and that we need to pay attention to the background conditions in which the supposed consent is (or is not) taking place. I want to focus on the second of these two considerations, but I shall say something briefly about the first.

It is a common objection to Kant's and Kantian ethics that intuitively permissible actions are prohibited. These fall within the domain of false negatives: actions that seem innocuous but are nevertheless not allowed if we follow the strict guidelines laid out by the Categorical Imperative (CI). For example, surprise birthday parties would be impermissible because they involve deception, and deception, even for a good end, cannot be made consistent with the CI.<sup>10</sup> The issue becomes more vexed if we interpret Kant's notorious murderer

maxim of deception or coercion can be agreed to or shared by those deceived or coerced. On the contrary, deception standardly works by revealing subsidiary intentions or aspects of action, which misleadingly point to some underlying maxim to which consent can be given. Deception only works when the underlying intention or proposal is kept obscure. The deceiver's actual maxim therefore cannot be consented to. A maxim of coercion does not have to be obscure—it may be brutally plain—but clearly denies victims the choice between consent and dissent.”

<sup>8</sup> *Ibid.*, p. 259.

<sup>9</sup> This marks a change (of which she is very well aware) from O'Neill's position in her earlier work. See, e.g., Onora Nell [O'Neill], *Acting on Principle*, New York: Columbia University Press, 1975.

<sup>10</sup> This example is taken from Barbara Herman, *The Practice of Moral Judgment*, Cambridge: Harvard University Press, 1993, chapter 7, p. 141.

at the door case as saying that it is impermissible to lie even to a homicidal maniac (or neoNazi) when s/he shows up at the door looking for the person you just helped to hide in your basement.<sup>11</sup> The task for Kantians becomes a sort of rescue attempt: how to rescue Kant from himself.

Now some Kant scholars argue that this is a misinterpretation of the murderer at the door case. For example, Wood tells us that the murderer at the door case involves the permissibility of making a lying declaration (analogous to perjury) rather than a simple lie.<sup>12</sup> If this is correct, then the murderer at the door case is not analogous to a case in which the Nazis knock on the door to ask whether you know the whereabouts of a Jew whom you happen to be hiding. Rather, the case is more like one in which you are asked to take the stand as a witness in a trial, and you know that if you tell the truth, there is a good chance that the defendant, whom you believe to be innocent, will be convicted and perhaps put to death—and the question is whether you should perjure yourself. For example, should Alexei Karamazov have given a lying testimony in *The Brothers Karamazov*? Unlike the Nazi case, the answer to this is not obviously “yes.”

Moreover, Kant’s *Lectures on Ethics* suggest that Kant himself thought that lying is sometimes permissible, and Wood points out that there are even conditions in which a lying declaration might be permissible in Kantian ethics.<sup>13</sup> I shall return to the murderer at the door case in the next section. The point is that O’Neill’s suggestion about fundamental principles is meant to apply exactly to cases in which deception is intuitively permissible even though it seems to be ruled out by the possible consent interpretation. The idea is that the deception involved in throwing a surprise party might be permissible (even on the possible consent interpretation) if it is not part of the person’s fundamental intentions—and the same could be said about lying to a Nazi.<sup>14</sup>

<sup>11</sup> Pallikkathayil seems to interpret the murderer at the door case in just this way. She points out that in the *Metaphysics of Morals*, Kant claimed that as far as right is concerned, people may say what they like, true or false, provided that it does not bear directly on matters of right (e.g., a lie about the terms of a contract). She then argues that in the essay in which he discusses the murderer at the door, Kant “does an about face regarding duties of right not to lie” (Pallikkathayil, *op. cit.*, p. 145). In other words, Pallikkathayil’s idea is that in the *Metaphysics of Morals*, there is no direct duty of right not to lie but in the essay discussing the murderer at the door, there is always a (direct) duty of right not to lie. But this is incorrect. The murderer at the door case is not discussing a duty of right not to lie (*simpliciter*) but rather a duty of right not to lie under specific conditions, and therefore it does not obviously give evidence of an about face on Kant’s part. I discuss this further in the text above.

<sup>12</sup> Allen Wood, *Kantian Ethics*, Cambridge: Cambridge University Press, 2007, chapter 11.

<sup>13</sup> *Ibidem*.

<sup>14</sup> O’Neill, 1985, *op. cit.*, p. 266: “In each case it is our fundamental proposals, principles, or basic intentions that must meet these conditions. We neither do nor can make it possible for others, even for others closely affected, to consent to or dissent from every aspect (or even every intentional aspect) of what we propose...”

This might strike some as somewhat sophistical: one winds up questioning when a deception really counts as a deception.<sup>15</sup> But I think this is probably an important question to raise.<sup>16</sup> However, I do not want to dwell on it here: I want to get to O'Neill's claims about understanding the background conditions for possible consent. As noted above, O'Neill's idea is that if we pay attention to the background conditions in which the supposed consent is (or is not) taking place, we shall see that the possible consent interpretation does not rule out some specific actions that seem to preclude consent.

The basic idea here is that there is more to consent than simply giving someone the opportunity to say "yay" or "nay."<sup>17</sup> If someone confronts you with

<sup>15</sup> *Ibid.*, p. 271: "Since it is only fundamental principles of actions (whether plans, proposals, policies, or intentions) that must meet these standards, superficial departure from them when acting on morally acceptable fundamental principles may be acceptable, or even required. *The jokes and surprises in which friendship may be expressed do not count as deceptions*; but if they were incident to action on other maxims might constitute fraud or serious disrespect or unacceptable paternalism" (my emphasis).

<sup>16</sup> This is a question that Kant himself seems to raise: "[i]f our untruth is in keeping with our main intent, then it is bad; but if I can avert a truly great evil only by this means, then... etc. Here goodness of heart takes the place of sincerity... A white lie is often a *contradictio in adjecto*; like pretended tipsiness, it is *untruth* that breaches no *obligation*, and is thus properly no lie. *Joking* lies, if they are not taken to be *true*, are not immoral..." (27, 62). (Following standard convention, I use the Academy pagination in my citations from Kant; the first number (27) refers to the volume in the Academy edition and the second number (62) refers to the page of that volume on which the citation occurs. All translations are taken from the Cambridge Guyer/Wood blue series translations.) See also 27, 700-702.

<sup>17</sup> Pallikkathayil distinguishes between the possible consent interpretation and the meaningful consent interpretation. She thinks that the possible consent interpretation requires one merely to give one's interlocutors the opportunity to signal consent or dissent (verbally or otherwise) regardless of whether one pays this signal any heed; the meaningful consent interpretation requires both that one give one's interlocutors the opportunity to signal consent or dissent and that one govern oneself in accordance with what one's interlocutors signal. An example will make this distinction clearer. If a murderer asks you whether s/he might shoot you and then does so despite your decided "no," then according to Pallikkathayil, the murderer has complied with the possible consent interpretation but not the meaningful consent interpretation (Pallikkathayil, *op. cit.*, section I.C., especially p. 127).

There are three problems with this, one rhetorical, one philosophical and one exegetical. The rhetorical problem with this is that Pallikkathayil introduces the meaningful consent interpretation only after dismissing the possible consent interpretation. But she dismisses the possible consent interpretation on the grounds that it is too strong (*i.e.*, it rules out permissible actions—it has false negatives). Given that the meaningful consent interpretation is *stronger* than the possible consent interpretation (*i.e.*, it rules out everything ruled out by the possible consent interpretation *and more*) it follows immediately (*a fortiori*) that the meaningful consent interpretation is too strong. So rhetorically, given Pallikkathayil's attack on the possible consent interpretation, no discussion of the meaningful consent interpretation should be needed beyond pointing out that it is stronger than the possible consent interpretation.

Philosophically, separating the possible consent interpretation from the meaningful consent interpretation is problematic because it vitiates Pallikkathayil's attack on the possible

a gun in a back alley and says, “your money or your life,” then the sense in which you “consent” to give that person your money when you hand it over is quite tenuous, indeed. The problem with the “consent” in this kind of case seems to lie in the background conditions (you have a gun to your head; the person “made you an offer you could not refuse”). This idea can be traced back to Marx, and it has been used also by Rawls, who combats Lockean historical accounts of justice with it, arguing that the basic structure of society must be regulated by the principles of justice as fairness to ensure the possibility of continued voluntary, just transactions.<sup>18</sup> But regardless of where one stands on these issues in political philosophy, O’Neill thinks that they are relevant for determining possible consent. This can be seen in the following two quotations:

...we need to understand what makes genuine consent to the more fundamental aspects of action possible. But there is no guarantee that any one set of requirements makes genuine consent possible in all circumstances. *There may be some necessary conditions, whose absence always makes genuine consent or dissent impossible, and other conditions which are needed to make consent possible only in some circumstances.*<sup>19</sup>

...coercion of less straightforward sorts may occur in some sexual relationships and transactions, including in relationships between prostitutes and their clients. Here the outward transaction may be an agreement between consenting adults. *But when we remember the institutional context of much (at least contemporary, western) prostitution, including the practices of pimping, brothel keeping, and various forms of social ostracism and consequent dependence on a harsh sub-culture, we may come to think that not all transactions between prostitutes and clients are uncoerced: but it may not be the client who coerces.*<sup>20</sup>

consent interpretation. That is, once we see how weak the possible consent interpretation is on Pallikkathayil’s account, it becomes unclear why it rules out all kinds of deception, force and coercion. The murderer described above seems to be engaged in a forceful action. But the murderer’s action would not be ruled impermissible by Pallikkathayil’s account of the possible consent interpretation.

Exegetically, the problem is that it is unclear whether anybody subscribes to this weaker version of the possible consent interpretation. For example, O’Neill makes it clear that she subscribes to what Pallikkathayil calls the meaningful consent interpretation: “if those closely affected have the possibility of dissent, they will be able to require an initiator of action either to modify the action, or to desist or to override he dissent. *But an initiator who presses on in the face of actively expressed dissent undercuts any genuine possibility of refusing the proposal and rather chooses to enforce it on others.* Any “consent” the proposal then receives will be spurious, and will not show that others have not been used, let alone that they have been treated as persons” (O’Neill, 1985, *op. cit.*, p. 259, my emphasis).

<sup>18</sup> John Rawls, *Justice as Fairness: a Restatement*, Cambridge: Harvard University Press, 2001, sections 14-16, p. 50ff, especially section 15.

<sup>19</sup> O’Neill, 1985, *op. cit.*, p. 259, my emphasis.

<sup>20</sup> *Ibid.*, pp. 268-269, my emphasis.

Both of these quotations reveal O'Neill's concern with background conditions. In the first quotation, O'Neill argues that there might be some background conditions that must be in place for consent to be possible in general, and there also might be some conditions that must be in place for consent to be possible in particular cases. In the second quotation, O'Neill gives an example of this. She argues that appreciation of the background conditions that lead some (perhaps many) to prostitution as a way of life might make us question whether the individual transactions in which they engage involve any real consent (despite their outward trappings). Perhaps consent was precluded long ago when, say, they were kidnapped and forced to sell their bodies. The issues here are complicated. Perhaps there is some sort of statute of limitations with regard to consent whereby past actions, if they are far back enough, do not infect the possibility of consent in future actions. It is far beyond the scope of this paper to discuss these issues much less to try to resolve them. So why bring this up?

The point of bringing this up is simple. As noted above, Pallikkathayil attacks the possible consent interpretation on the grounds that it rules out all actions that rely on deception, force and coercion for their nature despite the fact that some such actions are intuitively permissible (and condoned as such by Kant: the problem is thus both philosophical and exegetical). Pallikkathayil's examples are (1) the use of force in self-defense; (2) the use of force in compelling others to leave the state of nature and enter the civil condition; (3) the use of force on the part of the state in enforcing laws; and (4) the use of force on the part of the state in punishing criminals. But now we can see how O'Neill might respond to all of these examples.

With regard to the force used in self-defense, O'Neill might say that (1) the fundamental maxim of an agent acting in self-defense is not one of force and, therefore, not one that precludes consent and (2) maxims of self-defense plausibly are part of the background conditions that make consent possible. With regard to the force used in compelling others to leave the state of nature and enter the civil condition, O'Neill might point out again that (1) the fundamental maxim in this kind of case is not one of force and (2) it seems uncontroversial that the background conditions for consent are put in place by virtue of entering the civil condition (as noted above, for Rawls, part of the point of regulating the basic structure of society by the principles of justice as fairness is to ensure the continued possibility of meaningful consent). Finally, with regard to the use of force on the part of the state in enforcing laws and punishing criminals, again, O'Neill might point out that (1) the fundamental maxims of the agents in these cases are not maxims of force and (2) some sort of criminal justice system and executive branch of government form the background conditions required for consent to be possible.

The basic problem is that Pallikkathayil's attack on the possible consent interpretation begins with the claim that the possible consent interpretation

rules out all kinds of deception, force and coercion. But this claim is not true (*pace* Pallikkathayil): on O'Neill's account, the possible consent interpretation rules out deception, force and coercion only if they are part of an agent's fundamental maxim—and, further, some kinds of deception, force and coercion might be necessary as background conditions to make consent possible in the first place. Making consent possible on O'Neill's account is a tricky thing, far more nuanced than (say) giving someone the opportunity to give a nod of the head.

## *2. Korsgaard and possible consent*

But Pallikkathayil tells us that she is more concerned with Korsgaard's version of the possible consent interpretation than with O'Neill's.<sup>21</sup> So to be fair to Pallikkathayil, it is incumbent on us to examine also Korsgaard's view to see whether Pallikkathayil meets with more success there. Korsgaard, unlike O'Neill, would concede the first part of Pallikkathayil's attack on the possible consent interpretation: Korsgaard does think that the possible consent interpretation rules out all kinds of deception, force and coercion.<sup>22</sup> Further, Korsgaard thinks that because the possible consent interpretation rules out all kinds of deception, force and coercion, it rules out some permissible actions. In particular, Korsgaard reconstructs Kant's murderer at the door case, interpreting it as one involving the permissibility of a simple lie to save the

<sup>21</sup> Pallikkathayil claims, “[a]lthough O'Neill's and Korsgaard's views are very similar, I will focus on Korsgaard's view, since it is a bit more developed” (Pallikkathayil, *op. cit.*, p. 118). However, Pallikkathayil's reasoning here seems a bit shaky. For one thing, it does not seem to be the case that Korsgaard's version of the possible consent is more developed than O'Neill's. O'Neill's account of the possible consent interpretation occurs in the context of an article devoted entirely to figuring out what kinds of duty the formula of humanity prescribes, with space devoted both to actual consent and hypothetical consent; Korsgaard's main account of the possible consent interpretation occurs in the context of an article on whether Kant's ethics (in general) is able to take account of background conditions, and she discusses both the universality formulations and the kingdom of ends formulation of the categorical imperative. This is not by any means an objection to Korsgaard or the project she undertakes. The point is simply that Pallikkathayil's claim that Korsgaard's version of the possible consent interpretation is more developed than O'Neill's is *prima facie* implausible. Even more problematic is Pallikkathayil's claim that O'Neill's and Korsgaard's views are very similar: as noted in the previous section of this paper, O'Neill does not think that the possible consent interpretation rules out all use of force, deception and coercion; but as noted (below) in the present section of this paper, Korsgaard does think that the possible consent interpretation rules out all use of force, deception and coercion. Given that this is the focal point of Pallikkathayil's attack, it seems to me that O'Neill and Korsgaard are very dissimilar, indeed. I am sure Pallikkathayil has reasons, and perhaps good ones, for focusing on Korsgaard's view rather than O'Neill's. But the ones she cites give a misleading picture of O'Neill.

<sup>22</sup> Christine Korsgaard, *Creating the Kingdom of Ends*, Cambridge: Cambridge University Press, 1996, pp. 106-132, 137-140 and especially 295-296.



life of someone you are hiding from a murderer, and she argues that although the possible consent precludes deceit in this case, deceit is, in fact, permissible.<sup>23</sup> It seems like Pallikkathayil's two-pronged attack should apply: (1) this renders the possible consent interpretation exegetically implausible because it rules out actions that Kant thought were permissible and (2) this renders the possible consent interpretation philosophically implausible because it has false negatives.

Korsgaard's answer to the first prong of the attack is, I take it, quite simple: she is not engaged primarily in exegesis. She is trying to construct a plausible Kantian theory, a theory that has recognizably Kantian roots even if she ultimately reaches conclusions that are different from Kant's. I take it that Korsgaard would say that if Kant really thinks that FH condones some kinds of deception, force and coercion, then that is so much the worse for Kant. So the first prong of Pallikkathayil's attack is easily handled by Korsgaard. What about the second prong?

Korsgaard argues that although FH (by virtue of the possible consent interpretation) renders all deceit impermissible, which is counterintuitive, the so-called Formula of Universal Law (FUL) version of the CI does not. In particular, FUL gives the right answer to the murderer at the door case (as interpreted by Korsgaard): according to FUL, it is permissible to lie to a murderer at the door even though according to FH, it is not.<sup>24</sup> But according to Korsgaard, this discrepancy between the formulations of the CI does not indicate that one of them is "better" than the other. Rather, it indicates that they have different functions. Whereas FH is for ideal theory, FUL is for nonideal theory. The basic idea is that FH is the moral standard that we would use if everyone with whom we interacted had a good will. Perhaps we even can use FH when our immediate interlocutors do have (at least approximately) good wills. But even if not, we have FUL. FUL is there for the messy, everyday world in which we often have to deal with agents who have turned away from the good, so to speak—with people who act with malice aforethought, with people who are selfish, with people who are thoughtless, with people who are callous. In a word, with people who are nonideal. So Korsgaard has answers to both prongs of Pallikkathayil's attack.

However, Pallikkathayil does not stop there: she has two further attacks on Korsgaard's version of the possible consent interpretation. First, she tries to show that it does not work even for ideal theory.<sup>25</sup> That is, as we have seen, Korsgaard is committed to two claims: (1) FH tells us how we would be

<sup>23</sup> *Ibid.*, pp. 132-135.

<sup>24</sup> *Ibid.*, pp. 132-140. Korsgaard summarizes her findings as follows: "[i]f the foregoing casuistical analyses are correct, then applying the Formula of Universal Law and the Formula of Humanity leads to different answers in the case of lying to the murderer at the door. The former seems to say that this lie is permissible, but the latter says that coercion and deception are the most fundamental forms of wrongdoing" (*ibid.*, p. 143).

<sup>25</sup> Pallikkathayil, *op. cit.*, section I.C., especially p. 127.

supposed to behave if everyone with whom we interacted had a good will, and (2) FH rules out all actions that preclude consent (including all kinds of deception, force and conception). It follows immediately that Korsgaard is committed to (3) if everyone with whom we interacted had a good will, then all actions that preclude consent would be impermissible. But Pallikkathayil thinks this is false: Pallikkathayil argues that even in ideal theory, Korsgaard's version of the possible consent interpretation generates counterintuitive results.

Pallikkathayil's second attack is against the bare idea of FH as a standard of ideal theory. Pallikkathayil claims that by consigning FH to ideal theory, Korsgaard "undermines the status of the Formula of Humanity as a principle of practical reason."<sup>26</sup> In other words, Pallikkathayil seems to think that on Korsgaard's account, FH no longer can be considered as authoritative: it does not yield the right results to questions about what we may, may not or ought to do. So let us examine these attacks in more detail to see whether they really work.

The first attack is illustrated by an example in which we are to imagine ourselves looking out into the street at the passersby to see whether it is cold enough to wear a coat. One of the passersby signals us to stop, and Pallikkathayil claims that (1) on Korsgaard's interpretation of FH, we ought to respect the person's wishes and stop looking outside and (2) this reveals that Korsgaard's version of FH generates results that are counterintuitive even for an ideal theory.<sup>27</sup>

Now the idea behind (1) is that if we continue to look out the window even after the passerby has signaled us to stop, then our maxim is one that precludes consent and, thus, is impermissible according to Korsgaard's version of the possible consent interpretation of FH. That is, if someone (P) clearly indicates that s/he does not consent to an action in which someone else (Q) is engaged, then for Q to continue on is for Q to pay no heed to the fact that P does not consent (actively) to Q's action. But if Q pays no heed to the fact that P does not consent (actively) to Q's action, then Q is acting on a maxim that makes consent impossible. It follows immediately that in the example just described, Korsgaard's version of the possible consent interpretation of FH generates the conclusion that we ought to stop looking outside if one of the passersby signals us to stop.

The idea behind (2) is that "one may look out one's window at the activities on the street even if some particular person on the street objects."<sup>28</sup> Here Pallikkathayil calls upon the public/private distinction, a distinction that, according to Pallikkathayil, commits us to the idea that we are not always required to be sensitive to others' wishes about how and when they are observed—a distinction that (according to Pallikkathayil) should be affirmed even in ideal theory (Korsgaard notwithstanding).

<sup>26</sup> *Ibid.*, p. 125, n. 27.

<sup>27</sup> *Ibid.*, p. 127ff.

<sup>28</sup> *Ibid.*, p. 127.

But there are problems with both (1) and (2). First, it seems unlikely that the possible consent interpretation really would require us to stop looking outside if one of the passersby signals us to stop. If X's neighbor asks Y to stop looking over the fence at a party in X's backyard, the possible consent interpretation *might* require Y to stop looking over the fence and into X's backyard. But it would not require Y to stop looking over the fence at the sun. And in Pallikkathayil's example, the person who signals us to stop is on public property, so it seems like the strongest proscription that the possible consent interpretation could generate would be against looking *at that person*. Moreover, it should be remembered that an ideal theory applies to ideal conditions; presumably an ideal agent in ideal conditions would not mind if someone were looking out a window to determine whether it is cold enough to wear a coat. At first blush, this second remark might seem somewhat fatuous, but it is not: if the possible consent interpretation is part of ideal theory, then any putative counterexample to it must be realizable in an ideal world.

However, an even deeper problem with Pallikkathayil's example might arise with (2), with the claim that the example reveals that Korsgaard's version of FH generates results that are counterintuitive even for an ideal theory. That is, even if the possible consent interpretation generated the stronger proscription that Pallikkathayil attributes to it (it does not, but even if it did) it must be remembered that there is a distinction between having a right to X and the permissibility of X-ing. There are many things that are morally abhorrent but perfectly within our rights, and there are many things that are mildly distasteful but (again) well within our rights. Sometimes the rights that protect these activities are very important. But the point is that even if the possible consent interpretation generated the stronger proscription, this would be a result about the *permissibility* of looking out the window if one of the people being observed signals us to stop, not a result about whether looking out the window is *within our rights* if one of the people being observed signals us to stop—and this is not obviously counterintuitive for an ideal theory. Let me try to give a parallel case.

Suppose you are at a coffee shop and someone with a horrible disfigurement walks in. Now I think that there is something almost magnetic about such people. Like Leontius, we want to stare. And sometimes we do. Many of us even have had temporary (and/or permanent) problems that rendered us like this, drawing eyes like a sideshow barker but with this crucial difference: the attention is unsolicited. And it is unwanted.

Of course, I have a *right* to look at people like that. And that is a good thing: thank goodness I cannot be arrested for looking at someone slantwise. But looking away is not over and above what morality requires of me; it is simple respect, the barest recognition and sympathy for the feelings of another. So it seems to me that Pallikkathayil's first attack fails: she has not shown that Korsgaard's version of the possible consent interpretation generates counterintuitive results.

But what of Pallikkathayil's second attack? Recall that Pallikkathayil claims that by virtue of being consigned to ideal theory, FH loses its status as a principle of practical reason. Unfortunately, this attack is given only in a footnote, so Pallikkathayil does not explain exactly what she means. But if the idea is that a principle of practical reason is one that applies under all conditions (including nonideal conditions) then the objection is true but uninteresting. If the idea is that by virtue of being "relegated" to ideal theory, FH becomes useless or has no value—or even that other philosophers have not explored ideal theory—then the objection does not go through. Ideals play an important role in many areas of investigation, and an in depth discussion of them is beyond the scope of this paper. I note merely that one of the things that makes them useful is that the "calculations" are easier.

### *3. Conclusion*

Thus far I have defended the possible consent interpretation from Pallikkathayil. In particular, I have defended two very different versions of the possible consent interpretation from Pallikkathayil.<sup>29</sup> Whereas on O'Neill's account, the notion of possible consent does not rule out all kinds of deception, force and coercion, on Korsgaard's account, it does, but it applies to the domain of the ideal, so these proscriptions are not obviously counterintuitive.

But it should be remembered that the possible consent interpretation can tell us, at best, only what to do in order to avoid treating others as mere means. But that means that even at best, this interpretation of FH still has two serious exegetical problems: (1) FH tells us also never to treat ourselves as mere means, and duties to oneself (like the duty not to commit suicide from self-love) figure prominently in both the *Groundwork for a Metaphysics of Morals* and the *Metaphysics of Morals*. However, the notion of possible consent does not seem to apply to considerations of oneself: there does not seem to be any way to derive duties like the duty not to commit suicide from self-love from the possible consent interpretation. (2) In addition, FH tells us always to treat people as ends. The so-called "imperfect" duties that flow from this part of FH (like the duty to promote others' happiness) also figure prominently in both the *Groundwork for a Metaphysics of Morals* and the *Metaphysics of Morals*. However, acting on maxims to which it is possible for our interlocutors to consent seems a far cry from acting on maxims that involve treating these interlocutors positively as ends. Because the possible consent interpretation does not seem to be able to account either for duties to oneself or for imperfect duties, it can be at best only a piece of the larger puzzle of how Kant derives duties from FH. If one thinks that Kant is right to include such duties

<sup>29</sup> See my remarks in note 21 above.

in his ethical system, then these exegetical problems become philosophical problems, too.

However, proponents of the possible consent interpretation are aware of these two limitations, and they try to supplement the possible consent interpretation accordingly. Moreover, if the possible consent interpretation *can* account for all (even if only) our perfect duties to others, this is surely reason enough to take it quite seriously. Indeed, the notion of possible consent is not important only within the context of Kant's and Kantian ethics; the notion of possible consent is pervasive in law and ethics more generally and therefore independently interesting. As noted in section 1 of this article, possible consent is appealed to in core debates about distributive justice. It also shows up in discussions of criminal justice: many crimes, like rape, revolve around complicated questions about consent—whether it occurred (did that gesture really count as consent?), whether it was possible (was the victim drunk? How old was the victim?), etc—and similar questions come up in medical ethics when dealing with things like end of life issues, potential side-effects of a given course of treatment and the testing of new medicines, medical technologies and surgeries.

A thorough discussion of any (let alone all) of these issues is far beyond the scope of the present article. My goal has not been to show even that the possible consent interpretation succeeds in carving out the sphere of perfect duties to others let alone to apply the possible consent interpretation to any of these (well trod) areas of inquiry. My goal has been simply to show that the idea behind this interpretation should be taken seriously: if it falls, I do not think it will be the result of some easy take-down—even the interesting objections raised by Pallikkathayil seem to miss the mark.

## Bibliography

- Barbara Herman, *The Practice of Moral Judgment*, Cambridge: Harvard University Press, 1993.
- Immanuel Kant, *Lectures on Ethics*, ed. and trans. by Peter Heath, Cambridge: Cambridge University Press, 2001.
- Christine Korsgaard, *Creating the Kingdom of Ends*, Cambridge: Cambridge University Press, 1996.
- Japa Pallikkathayil, "Deriving Morality from Politics: Rethinking the Formula of Humanity," *Ethics* 121.1 (2010), pp. 116-147.
- Onora Nell [O'Neill], *Acting on Principle*, New York: Columbia University Press, 1975.
- Onora O'Neill, "Between Consenting Adults," *Philosophy and Public Affairs* 14.3 (1985), pp. 252-277.
- John Rawls, *Justice as Fairness: a Restatement*, Cambridge: Harvard University Press, 2001.
- Allen Wood, *Kantian Ethics*, Cambridge: Cambridge University Press, 2007.