



Paternalism, Respect, and the Will

Author(s): Daniel Groll

Reviewed work(s):

Source: *Ethics*, Vol. 122, No. 4 (July 2012), pp. 692-720

Published by: [The University of Chicago Press](#)

Stable URL: <http://www.jstor.org/stable/10.1086/666500>

Accessed: 27/08/2012 09:49

Your use of the JSTOR archive indicates your acceptance of the Terms & Conditions of Use, available at <http://www.jstor.org/page/info/about/policies/terms.jsp>

JSTOR is a not-for-profit service that helps scholars, researchers, and students discover, use, and build upon a wide range of content in a trusted digital archive. We use information technology and tools to increase productivity and facilitate new forms of scholarship. For more information about JSTOR, please contact support@jstor.org.



The University of Chicago Press is collaborating with JSTOR to digitize, preserve and extend access to *Ethics*.

Paternalism, Respect, and the Will*

Daniel Groll

In general, we think that when it comes to the good of another, we respect that person's will by acting in accordance with what he wills because he wills it. I argue that this is not necessarily true. When it comes to the good of another person, it is possible to disrespect that person's will while acting in accordance with what he wills because he wills it. Seeing how this is so, I argue, enables us to clarify the distinct roles that the wills of competent and incompetent people should play in third-party deliberations about their welfare.

Bob is in hospital after a stroke. His doctor tells him that he requires a percutaneous endoscopic gastrostomy (a PEG) in order to receive enough nutrition to survive.¹ He understands what the doctor tells him about the procedure (the minimal risks and the substantial benefits). In other words, he is competent. But he refuses to undergo the surgical procedure simply because he does not think the surgery is worth the risk. He explains this to the doctor and tells her that he would rather do nothing than undergo the procedure.

I assume everyone agrees about what Bob's doctor should do. The decision primarily concerns Bob and his health, so the doctor ought to respect his will and not perform the surgery, even if she thinks it would be good for Bob to have it.² In general, we think that when it comes to what is

* Many thanks to Tom Beauchamp, Jason Bridges, Daniel Brudney, Emily Carroll, David Finkelstein, Maggie Little, Jennifer Lockhart, Thomas Lockhart, Micah Lott, Dorothy MacKinnon, Martha Nussbaum, two anonymous referees from *Ethics*, and a series of associate editors at *Ethics* for very helpful comments on earlier drafts. Jason Decker and Alice MacLachlan are owed special thanks for helping me significantly retool the essay in light of the comments from referees. Finally, my greatest thanks goes to Connie Rosati: the essay has improved immeasurably as a result of her very patient and helpful comments.

1. A PEG is "a surgical procedure for placing a feeding tube without having to perform an open laparotomy (operation on the abdomen). The aim of a PEG is to feed those who cannot swallow." This definition is from MedicineNet.com, "Definition of PEG," 2011, <http://www.medterms.com/script/main/art.asp?articlekey=6692>.

2. This is consistent with thinking that she ought to try to persuade Bob to change his mind.

Ethics 122 (July 2012): 692–720

© 2012 by The University of Chicago. All rights reserved. 0014-1704/2012/12204-00XX\$10.00

good for another person who is “in the maturity of [his] faculties,” it is objectionably paternalistic to not follow that person’s will.³ When it comes to what is good for another person, we respect that person’s will by acting in accordance with what he wills because he wills it. As Derek Parfit puts it, “In helping other people to fulfill or achieve their desires or aims, we respect these people’s autonomy, and avoid paternalism.”⁴

I argue that this is not necessarily true. When it comes to what is good for another person, it is possible to disrespect that person’s will by acting in accordance with what he wills because he wills it. This no doubt sounds counterintuitive. Once we get clear, however, on two distinct ways in which we can act in accordance with another’s will because it is that person’s will, there is good reason to accept my claim. In some cases, I will argue, treating someone’s will as decisive in one way, but not another, counts as disrespecting that person’s will.

As should be clear, my central interest is in the relationship between a first person’s actions and a second person’s will, where the first person acts for the good of the second person. This relationship is at the heart of discussions about paternalism. My aim, however, is not to offer a complete analysis of paternalism, let alone a defense of that analysis. This is because I simply take for granted various other ideas that would need to be fleshed out in a complete analysis of paternalism. So, for example, I take for granted that acting paternalistically involves acting for the good of the paternalized subject. But some reject this.⁵

My focus, instead, is on one aspect of paternalistic action—namely, the relationship between the paternalist’s action and the paternalized subject’s will. I want to show the shortcomings of the dominant way of conceptualizing this relationship (which I discuss presently) and propose an alternative.

Having said that, I refer throughout to the kinds of cases I’m interested in as instances of paternalism and also, in discussing the feature of paternalism I am paying close attention to, talk in terms of

3. J. S. Mill, *On Liberty* (Indianapolis: Hackett, 1978). Some might think that “objectionably paternalistic” is pleonastic on the grounds that paternalism is, by definition, objectionable. I disagree, although I think it is presumptively—by which I mean *prima facie*—objectionable. I discuss this below, in Sec. III. I would add that, of course, one might not follow the will of another for reasons that do not have to do with what is good for that person: perhaps following the person’s will would be too onerous on oneself or others.

4. Derek Parfit, *On What Matters* (Oxford: Oxford University Press, 2011), 1:66.

5. For example, Shiffrin claims that “traditional accounts of the paternalist motive misstep by insisting that the paternalist be motivated by concern for the interests or welfare, more broadly construed, of the paternalistically treated agent.” Instead, Shiffrin argues, “Behavior may be paternalist if the motive behind it is simply that the (putative) paternalist knows better than the agent, or may better implement, what the agent has authority for doing herself.” See S. V. Shiffrin, “Paternalism, Unconscionability Doctrine, and Accommodation,” *Philosophy and Public Affairs* 29 (2000): 205–50, 215, 216.

figuring out what paternalism amounts to. But nothing much hangs on the word. I am not trying to convince friends of other conceptions of paternalism (about which I say more presently) that they ought to use the word differently than they do: if they insist that the phenomenon I identify below is an instance of disrespecting another's will but not of paternalism, then so be it.

I liberally use the term "paternalism" in what follows, however, for two reasons. First, it is convenient. Second, it helps shed light on an area where questions of paternalism are front and center—namely, clinical medicine. The distinction I make between two ways of treating another's will as decisive allows us to clarify—in a way that standard accounts of paternalism do not—the distinct roles that the wills of competent and incompetent patients can play in third-party deliberations about their welfare.

These roles are not well understood. At the medical ethics case conferences I used to attend regularly, it was not uncommon to hear what I take to be two mistaken views on this matter. On the one hand, some clinicians thought that if a patient is deemed incompetent, then his current wishes count for nothing qua expressions of what the person wants.⁶ Others, on the other hand, seemed to think that respecting the patient's will (or respecting his autonomy, as it is sometimes put) was still centrally important in determining what to do, even though proponents of this idea admitted that the patient was incompetent. I take each position to be motivated by what it sees as the unreasonableness of the other: it is surely wrong to think that an incompetent patient's will counts for nothing, and it is also wrong to think that respecting the patient's will in these cases is just the same as respecting a competent patient's will. Getting clear on the different ways we can treat another's will as decisive, I argue, allows us to avoid both answers: the incompetent patient's will counts for something, but it plays a different "logical role," so to speak, than a competent patient's will in determining what to do.⁷

6. The "current" is crucial, since everyone agrees that wishes expressed while competent about what one would want in situations where one is not competent are highly relevant to making the right decision.

7. The question of how to understand the wishes of incompetent patients has received some attention from philosophers who, unsurprisingly, avoid the extreme views I articulated above. Ronald Dworkin distinguishes between "critical" and "experiential" interests and argues that the former—which only a competent person can have (since it is partly determinative of being competent that one has critical interests)—trump the latter (Ronald Dworkin, *Life's Dominion: An Argument about Abortion, Euthanasia, and Individual Freedom* [New York: Knopf, 1993], 203–16). So, the wishes of an incompetent patient will count for something qua wishes but not in the same way that the wishes of a competent patient count. But we need to be careful when we talk of someone being incompetent, since the real question is whether the person is incompetent with respect to decision, or issue. X. Agnieszka Jaworska, in responding to Dworkin, makes this point while con-

For reasons I explain below (Sec. III), this point is obscured, or not noticed at all, if we work with standard conceptions of paternalism. This is because standard conceptions of paternalism assume that to act paternalistically one must act against the will of the other:

1. "Paternalism is the interference of a state or an individual with another person, against their will, and justified by a claim that the person interfered with will be better off or protected from harm."⁸
2. "Paternalistic policies are restrictions on a person's liberty which are justified exclusively by consideration for that person's own

vincingly arguing that in many cases we ought to respect the current wishes of (many) patients suffering from Alzheimer's because they are still competent with respect to many issues related to their basic well-being. Consequently, Jaworska concludes, they "still have *authority* concerning their well-being" (emphasis added; Agnieszka Jaworska, "Respecting the Margins of Agency: Alzheimer's Patients and the Capacity to Value," *Philosophy and Public Affairs* 28 [1999]: 109, 105–38). Rebecca Dresser also disagrees with Dworkin, but on different grounds. She buys Dworkin's distinction and accepts that people with advanced dementia may only have experiential interests. Even so, she argues, there may well be cases where we ought to set aside the advance directive of a patient currently suffering from dementia if the patient's current life offers a reasonable degree of "pleasures and satisfactions" (Rebecca Dresser, "Dworkin on Dementia: Elegant Theory, Questionable Policy," *Hastings Center Report* 25 [1995]: 36, 32–38). Dresser does not, however, appear to accord any independent weight to the wishes of incompetent patients. My discussion below does not address the question of how we should deal with conflicts between a person's previously and currently expressed wishes when the cognitive abilities of the person have declined significantly from then to now. Even so, my account of the proper place of an incompetent patient's wishes in determining a course of treatment is not unrelated to Dworkin's discussion inasmuch as I, like Dworkin, think that the wishes of the incompetent patient count but not in the same way, or to the same extent as, the wishes of competent patient. The canonical discussion of how to make decisions for others in a medical context is found in A. E. Buchanan and D. W. Brock, *Deciding for Others: The Ethics of Surrogate Decision Making* (Cambridge, MA: Cambridge University Press, 1989), although Brock and Buchanan do not take up the precise issue I deal with here. I say a little more about how my view relates to some of what they say below (Sec. III).

8. Gerald Dworkin, "Paternalism," in *The Stanford Encyclopedia of Philosophy*, ed. Edward N. Zalta (Stanford, CA: Stanford University Press, 2010). The idea that the paternalist acts against the will of the paternalized subject is clearly part of Dworkin's understanding of paternalism even though it does not figure in his formal definition of paternalism. It does, however, show up in other places. When discussing when, if ever, paternalism is justified, Dworkin puts the (plausible) options like this: "Either we are never permitted to do good for others against their wishes, and in ways which limit their liberty, or we are permitted to do so" (Dworkin, "Paternalism"). And in an earlier work he claims that "It begins to look as if the only condition that will work [in defining paternalism] is one that depends upon the fact that the person who is being treated paternalistically does not wish to be treated that way" (Gerald Dworkin, *The Theory and Practice of Autonomy* [Cambridge: Cambridge University Press, 1988], 123). Dworkin's way of putting the point here makes it sound like he rejects this conclusion. But he does not; it is just that he thinks that this condition, while necessary, is not sufficient to define an action as paternalistic.

good or welfare, and which are carried out either against his present will . . . or against his prior commitment.”⁹

3. “[Paternalism] . . . consists in treating adults as if they were children, or older children as if they were younger children, by forcing them to act or forbear in certain ways . . . ‘for their own good,’ whatever their wishes in the matter.”¹⁰
4. “P behaves paternalistically towards Q [if and only if: . . .] P discounts Q’s belief that P’s behavior does not promote Q’s good.”¹¹
5. “Paternalism is the intentional overriding of one person’s autonomous choices or actions by another person, where the person who overrides justifies the action by appeal to the goal of benefitting or of preventing or mitigating harm to the person whose choices or actions are overridden.”¹²

All of these conceptions of paternalism contain what I will call the *CONTRARY TO WILL CRITERION*, some more explicitly than others. Dworkin and Arneson explicitly say that to act paternalistically is to act against the will of the paternalized subject. And to override someone’s choice or action (Beauchamp) or to force someone to act or forbear (Feinberg) is, clearly, to act against her will. The *CONTRARY TO WILL CRITERION* is better hidden in Archard’s account, but it is still there: what P discounts or ignores is Q’s belief that P’s action is not good for Q (and where Q is after his own good).¹³ So, insofar as P acts in a way that Q thinks is not for his (Q’s) own good, P must act contrary to Q’s will, since both P and Q are after Q’s good.

I want to show that inasmuch as our understanding of paternalism focuses on the *CONTRARY TO WILL CRITERION*, we risk losing sight of—or never getting into view—the morally important different ways that one can comply with the will of another because it is that person’s will. There is, I argue, another criterion—what I call the *DECISIVE REASON CRITERION*—that allows us to capture these differences.

I build my case for this alternative criterion as follows: In Section I, I introduce two problems for the *CONTRARY TO WILL CRITERION* that give us reason to look for an alternative and also point to what that alternative might be—namely, that what matters when deciding whether someone has acted paternalistically is whether she has treated the will of the other person as decisive in determining how she should act. In Section II, I draw

9. R. J. Arneson, “Mill versus Paternalism,” *Ethics* 90 (1980): 470–89, 471.

10. Feinberg actually calls this “presumptively blamable benevolent paternalism.” See Joel Feinberg, *Harm to Self* (Oxford: Oxford University Press, 1989), 5.

11. David Archard, “Paternalism Defined,” *Analysis* 50 (1990): 36–42, 36.

12. T. L. Beauchamp, *Standing on Principles: Collected Essays* (Oxford: Oxford University Press, 2010), 103.

13. Archard, “Paternalism Defined,” 39.

on the work of Joseph Raz and Stephen Darwall to show that there are two ways in which we can treat the will of another as relevant: we can treat it as either *structurally* decisive or *substantially* decisive. As indicated above, I show how this distinction allows us to understand the proper role of the will of an incompetent person in deliberations about his good. In Section III, I use the distinction between substantial decisiveness and structural decisiveness to arrive at the DECISIVE REASON CRITERION of paternalism. With the DECISIVE REASON CRITERION on board, the central problem with the CONTRARY TO WILL CRITERION comes into sharp focus: it does not allow us to see, in a way that the DECISIVE REASON CRITERION does, how we can fail to accord proper respect to another's will in doing what he wants because he wants it. But, for reasons that become clear below, not all cases of acting paternalistically (on my conception of paternalism) count as disrespecting the paternalized subject's will. In Section IV, I offer two accounts of what makes it the case that paternalism (given my conception of it) is impermissible, when it is: the AUTHORITY account and the SOUND JUDGMENT account. The AUTHORITY account seems like the natural one in many contexts, but I will argue that there are problems with it that lead us to the SOUND JUDGMENT account.

I. THE CONTRARY TO WILL CRITERION: SOME PROBLEMS

Let's return to Bob. Imagine that his doctor, who doesn't think much of the strong antipaternalist cast of modern clinical practice,¹⁴ decides to perform the surgery anyway. Here we have a paradigm case of someone acting paternalistically. It is worth noting that the doctor's paternalistic decision is consistent with her taking very seriously Bob's desire not to have the surgery. We can imagine the doctor weighing Bob's will alongside the other considerations and deciding to do the surgery despite Bob's will that it not be done. This shows that the paternalist need not be insensitive or indifferent to the will of the paternalized subject. The doctor does not ignore or discount Bob's will, nor does she treat it as irrelevant. It just turns out that it was not a sufficiently weighty consideration to stop her from deciding to operate.

The CONTRARY TO WILL CRITERION can explain, at least in part, why the doctor acts paternalistically here: Bob's physician acts contrary to Bob's will in performing the surgery. Even so, I want to convince you that this explanation puts the emphasis in the wrong place and that the CONTRARY TO WILL CRITERION—despite its apparent clear application in this case—has problems as part of a general criterion of paternalism.

First, by putting emphasis on whether someone has, in fact, acted contrary to the will of the other, the CONTRARY TO WILL CRITERION would

14. And isn't afraid of the law!

seem to have nothing to say about cases where the other person has yet to have a will with respect to the issue at hand. Let's call this the No Will problem. Suppose that Bob does not yet have a view about whether he should have a PEG because he does not yet know he needs a PEG. The physician cannot avoid acting paternalistically, on the grounds that she has not acted contrary to Bob's will, by keeping Bob in the dark and doing the surgery without his knowledge!

More realistically, we might imagine a case where someone who is in debt, let's call him Ernest, is, unbeknownst to him, having his debt paid down by his friend Sabina. Were Ernest to find out, he would be very upset. But as it has never even occurred to him that anyone would pay down his debts for him, he has no will with respect to the issue. He hasn't even given thought to how he would feel if he found out someone was helping him in the way Sabina is. Sabina, then, is not acting contrary to Ernest's will. Even so, she surely acts paternalistically, particularly if she is also actively working to keep her increasingly less indebted friend in the dark about her generous activities.¹⁵

Second, the CONTRARY TO WILL CRITERION seems to get things wrong in cases where a person's actions, intended for the benefit of the other, accidentally accord with the other's will. Let's call this the Accidental Concordance problem. Suppose, for example, that Ernest desperately wants Sabina to help pay off his debt, but he simply hasn't gotten around to asking Sabina for help. Sabina knows that Ernest is deeply in debt and decides to help him by paying it down. But, crucially, she does so without any thought to whether this is what Ernest would want her to do (she's that kind of person).¹⁶ She does not act contrary to Ernest's will here. Quite the opposite. But her acting in accord with his will is a matter of luck. Now, perhaps not everyone would want to say that Sabina acts paternalistically here. Perhaps we have a case of a paternalistic intention but not actual paternalism. It is, however, uncontroversial that Sabina

15. Proponents of the CONTRARY TO WILL CRITERION might say that this problem can be avoided by saying that one acts paternalistically in acting against another's will or against what the other's will would be (although this might threaten to make too many actions paternalistic). In other words, the person's will should be understood dispositionally. So, on this view, Bob's doctor acts paternalistically because Bob would not consent to the surgery if asked. As I say in just a moment, whether the CONTRARY TO WILL CRITERION fully fleshed out succumbs to this objection doesn't matter much for my purposes since it faces the more central problem I highlighted in the introduction and which I discuss below (Sec. III). My purpose here is just to motivate the search for, and build up to, my alternative criterion by showing that the simple articulation of the CONTRARY TO WILL CRITERION has problems.

16. Importantly, it is not that she assumes she has something like standing permission to pay off his debts or that this is what he really wants. I say a little more about this qualification below (n. 37) in relation to standard cases of gift giving.

here evinces a paternalistic attitude. And the CONTRARY TO WILL CRITERION does not tell us why.

A friend of the CONTRARY TO WILL CRITERION might claim that it does. For it is natural to think that while Ernest does want Sabina to pay off his debts, he wants her to do so only as a result of Ernest giving her permission to do so. By paying off his debts without a thought to what Ernest wants, she acts contrary to his will. But this doesn't solve the problem of Accidental Concordance. For we could imagine that what Ernest really wants is for Sabina to pay off his debts without giving his will any thought (he is too embarrassed to ask her and would simply like the problem to go away by her taking care of it without consulting him or considering what he wants). And this is what she does. Her action accords with Ernest's will perfectly, but accidentally, since she does not give any thought to what Ernest wants in this situation. Consequently, her action (at least) evinces a paternalistic attitude. And, once again, the CONTRARY TO WILL CRITERION does not tell us why.

I do not pretend that either the No Will problem or the Accidental Concordance problem as I have briefly presented them constitute a knock-down argument against the CONTRARY TO WILL CRITERION. Perhaps proponents of the criterion will be able to get around both problems with various modifications to the criterion. But the problems at least give us some reason to look for another account of the relationship between the doctor's or Sabina's actions and Bob's or Ernest's wills, which explains why the doctor and Sabina fail to respect Bob and Ernest. This is particularly true when we focus on the Accidental Concordance problem, for it is there that we find the seed of an alternative account. The problem there was not that Sabina acted against Ernest's will, for she acted in accord with it. The problem is that her action did not accord with it because of, or in response to, Ernest's will. Put another way, the problem with Sabina and the doctor is that they did not treat Ernest's and Bob's wills as *decisive* in determining what to do: that Bob didn't want the surgery or that Ernest did want help paying down his debt didn't determine the doctor's or Sabina's course of action (even if, in the latter case, Sabina did exactly what Ernest wanted). It is this idea that gets to the heart of what is problematic in the above cases.

II. STRUCTURAL DECISIVENESS VERSUS SUBSTANTIVE RELEVANCE

The idea on the table is this: central to understanding the above cases as instances of one person failing to respect the will of another (when acting for that person's good) is not that one person has acted contrary to the will of the other person; rather, it is that the acting agent has not treated the other person's will as decisive in determining what she should do.

But what does it mean to treat the will of another as decisive? The question does not have a single answer because there are (at least) two things we might mean. Or, to put the point slightly differently, there are two ways in which we might treat the will of another as decisive. We can get at them by introducing Carl to our stable of unfortunate people in need of a PEG:

Carl and the PEG: Carl, like Bob, requires a PEG. Unlike Bob, however, Carl is deemed incompetent: in addition to being somewhat confused about where he is and why, he does not have an adequate grasp of the risks and benefits of the surgery. He also has extremely conspiratorial beliefs about his doctor. He understands, however, that in order to get the PEG he will require surgery. Carl absolutely insists that he does not want the PEG because he does not want surgery. Moreover, Carl has never been counted as a competent decision maker, so there is no advance directive or anything of the sort that explains what a competent Carl would will in such a situation.¹⁷

Let's suppose that leaving aside Carl's will about getting a PEG, it is in his best interest to get it.¹⁸ The question I want to ask is this: What role should Carl's will play in determining what to do? Remember that Bob's case is (relatively) easy. He insists he does not want a PEG, and he is competent to make that decision. So, he should not get the PEG unless he changes his mind. Carl's case is less easy. Like Bob, he needs the PEG and very clearly does not want it. But, unlike Bob, he is not a competent decision maker. So, it is an open question whether he ought to get the PEG over his protests. Suppose, however, that Carl's surrogate ultimately decides not to give Carl the PEG because he so clearly does not want it. Suppose, further, that Carl not wanting it weighs more heavily in the surrogate's deliberation than anything else and, indeed, more than the combination of all the other considerations.

Now notice this: in both cases, the well-behaved doctor and Carl's surrogate will cite the will of Bob and Carl (respectively) in explaining her decision: "Bob/Carl really needs a PEG, but I won't give him one because Bob/Carl insists he does not want it." In this sense, both men's wills are decisive in determining the doctor's/surrogate's course of action.

But putting things like this obscures an important difference between the two cases. When Bob declares that he does not want surgery, his will is authoritative. This means that Bob, and in this case no one else, is the *de jure* ultimate decision maker in Bob's case. In other words, Bob's

17. Should we think of Carl as even having a will? Surely we should. Even if Carl's will is diminished or defective, it still makes sense to see him as willing various things in a way that someone in a coma cannot and as willing things in a more robust way than a very small child or an animal can be said to will things.

18. I come back to this important qualification presently.

will grounds a legitimate demand that the surgery not be performed; he is effectively issuing an order that he not have the surgery, an order that only he is authorized to give. And it is in the conceptual nature of an order that it be treated as what I will call structurally decisive in determining what to do—it is meant to supplant the reason-giving force of other considerations not because it outweighs those other considerations but because it is meant to silence, or exclude, those other considerations from the practical deliberations of the subject of the demand, in this case Bob's doctor. To paraphrase R. Dworkin, Bob's demand is meant to trump whatever other considerations are on the table.¹⁹

We can put this idea as follows: the force of the reason not to do the surgery that is grounded in Bob's demand is insensitive to considerations of Bob's good. This means that the doctor cannot discount Bob's demand by pointing out that following it is not good for Bob. Relatedly, the normative force of Bob's demand is not properly assessed by determining what good (for Bob) comes from following it.²⁰ The point of an authoritative demand in this context is to render such appeals to what is good

19. Ronald Dworkin, "Rights as Trumps," in *Theories of Rights*, ed. Jeremy Waldron (Oxford: Oxford University Press, 1984), 153–67. My idea here is clearly indebted to the idea of an "exclusionary reason" developed by Raz (Joseph Raz, *Practical Reason and Norms* [Oxford: Oxford University Press, 1999]) and the more recent idea of a "second-personal reason" developed by Darwall (Stephen L. Darwall, *The Second-Person Standpoint: Morality, Respect, and Accountability* [Cambridge, MA: Harvard University Press, 2006]). Both ideas basically amount to the claim that some practical reasons function to silence, or preempt, other considerations, and this is the idea behind my conception of someone's will being structurally decisive. But neither thinker has, to my knowledge, deployed the idea of a special kind of authoritative practical reason to shed light on the issue of paternalism or how the wills of people that are not competent (like Carl) should figure into our deliberations about how we should treat them (for their own good). Nor have they discussed, as I do below, the conditions under which it is permissible to not treat the will of another as structurally decisive for that person's own good.

For a discussion of the connection between Darwall's idea and Raz's, see Stephen L. Darwall, "Authority and Second-Personal Reasons for Acting," in *Reasons for Action*, ed. David Sobel and Steven Wall (Cambridge: Cambridge University Press, 2009), and "Authority and Reasons: Exclusionary and Second-Personal," *Ethics* 120 (2010): 257–78. In the first article, Darwall argues that Raz's conception of when one person has practical authority over another will not work because, in brief, it would seem to create authority relations where there are, in fact, none. In the second article, he argues that Raz's account of the conditions under which exclusionary reasons, and so authority relations, are generated will not work since exclusionary reasons are generated only where an antecedent authority relation already exists. The real issue between Raz and Darwall, then, is about whether authority relations can, ultimately, be reductively explained: Raz thinks they can, while Darwall thinks they cannot.

20. This leaves open the question of whether we can assess Bob's right or authority to make such demands in terms of some good that is promoted or protected by giving people like Bob (i.e., competent patients) such authority. In other words, one might think that the practice of making authoritative demands in all its guises is ultimately explained in terms of the good. Some version of this thought is developed and defended by Mill, *On Liberty*; John Rawls, "Two Concepts of Rules," *Philosophical Review* 64 (1955): 3–32; Joseph Raz, *The*

for Bob, at least on the part of the doctor, irrelevant. Notice the disanalogy here with what we might call theoretical authority, that is, authority concerning what is true: the reason-giving force of an authoritative theoretical claim (i.e., its ground as a reason for belief) necessarily varies with the force of the independent evidence in support of the claim in question. The more independent reasons we have for thinking the claim is false, the less reason we have to believe a purportedly authoritative claim to the contrary. To take the extreme case, it makes no sense to say, "I know that claim A is false, but your claim that it is true gives me an independent reason to believe it." But it makes perfect sense to say, "I know that demand A will result in no good, but your demand that A nonetheless gives me an independent and perhaps sufficient reason to act on it." What this shows is that the reason-giving force of particular demands (if not the actual practice of making authoritative demands) is not measured in terms of their ability to track the good (whereas the reason-giving force of particular authoritative theoretical claims is measured in terms of their ability to track the true: there is no other reason-giving dimension on the theoretical side²¹). Bob's will, then, grounds a reason to act that, at least from the point of view of the doctor, is normatively forceful apart from considerations of Bob's good.

The situation with Carl is very different. Since Carl is incompetent, he cannot make authoritative demands. And this means that Carl is not the *de jure* ultimate decision maker with regard to the decision to have surgery—someone else is the proper decision maker. So now we can ask, what factors should go into the surrogate decision-maker's deliberation? And more specifically, how might Carl's will figure in that deliberation, given that he has never been competent (and so never made clear his wishes while competent)?²² It might seem that, as with Bob, we face a conflict between two kinds of reasons: those based on Carl's good (which tell in favor of getting the PEG) and those based on his will (which tell in favor of not getting the PEG). But this cannot be right for the simple reason that Carl's will cannot ground authoritative demands. This means that it cannot ground a practical reason the force of which is insensitive to considerations of what is good. So, whatever reason-giving force Carl's will

Morality of Freedom (Oxford: Oxford University Press, 1988); and Richard Kraut, *What Is Good and Why: The Ethics of Well-Being* (Cambridge, MA: Harvard University Press, 2009). Critics of this idea include Feinberg (*Harm to Self*) and Darwall (*The Second-Person Standpoint*). Nothing I say depends on taking a side on this issue.

21. This suggests that even if there are second-personal theoretical reasons, they are importantly different than second-personal practical reasons. For discussion of the possibility of second-personal epistemic reasons, see Benjamin McMyler, *Testimony, Trust, and Authority* (Oxford: Oxford University Press, 2011).

22. Had he done so, the surrogate decision maker is subject to competent-Carl's authoritative demand, thereby rendering Carl and Bob's situations effectively the same, although on this point see above, n. 7.

has in this situation must relate to the question, what is good for Carl?²³ Unlike with Bob, where in addition to considering his good we must also consider his right (i.e., his right to make claims on those taking care of him), the only thing that we can consider in Carl's case is Carl's good. This is because Carl's will, unlike Bob's, cannot ground authoritative demands. There are no considerations of Carl's right as determined by his will.²⁴

Does this mean that Carl's will—his expressed wish not to have surgery—can play no role in the surrogate's deliberation? Not at all. It is just that Carl's will to not have surgery is one factor among others that goes into an all-things-considered judgment about what is good for Carl. In other words, the relevant question is whether it is good (for Carl) to follow his will in this situation. It might turn out that what is best for Carl is letting him have his way because it is his way.²⁵

But how could this be, given that he is incompetent? Following an incompetent patient's will can contribute to his good derivatively inasmuch as following his will causally leads to greater overall well-being. So, for example, it might be that given that an incompetent patient does not want a procedure performed, the psychological and physical trauma for the patient of actually performing the procedure outweighs the intended benefit. This is especially true with procedures or treatments that require ongoing compliance (dialysis, for example). In many cases the cure may be worse than the disease precisely because the patient's strong resistance to the procedure or treatment, if carried out, will lead to deleterious effects on other aspects of the patient's well-being. Perhaps this is the case with Carl.²⁶ I say that following his will here contributes to his good only derivatively because following his will is not accorded any weight as a constituent of his well-being: the value of following his will derives entirely from its connection to promoting his well-being, where that is construed as not being constituted, even in part, by Carl having his will followed.

The derivative link between following an incompetent patient's will and his well-being is all I need to make my point that what might be best

23. Assuming that we are not concerned with anyone else's good here.

24. There are, of course, considerations of Carl's right that stem from the fact that he is a patient and a person (i.e., all patients and persons have certain rights), and the surrogate must consider these. The point is just that Carl's will to not get the PEG grounds no such right.

25. This point is not just relevant for Carl. It may also be true that part of Bob's good consists of his being self-determining. The key point is just that whatever reason the good of being self-determining gives us to respect Bob's and Carl's wills, Bob's ability to issue an authoritative demand gives us a different kind of reason, one that is not grounded in considerations of Bob's good.

26. Although it might not be if his life depends on getting the PEG.

for Carl is letting him have his way because it is his way. But I want to suggest that there is a second, and less obvious, way the incompetent patient's will can contribute to his good: it can do so nonderivatively, that is, simply in virtue of being followed. We are inclined to think that this is true when dealing with competent patients and people in general: part of what it is for your life to go well is for you be autonomous. And part of what it is to be autonomous is to have your will play a central role in determining the shape of your life. So this point clearly applies to Bob: quite apart from his authority to refuse the PEG, it is plausible to think that part of his good consists in being self-determining, which in this case involves not getting the PEG.²⁷

But can we extend this idea to incompetent patients like Carl? One might think not. The idea that it is part of one's well-being to have one's will play a central role in determining the shape of one's life isn't true for just anyone; instead, it applies to people that have rational or autonomous wills.²⁸ And incompetent patients don't have those—if they did, they would be medically competent. So there is not, in their case, any nonderivative contribution to their good that obtains simply in virtue of following their wills.

I want to resist this thought. Naturally, much depends on what it means to "have a will." Some incompetent patients cannot plausibly be said to have wills at all. Others perhaps do, but they are sufficiently distorted by mental illness or some other affliction that we accord following them no nonderivative value. Even so, I think it is possible for someone to fall short of the threshold required for medical competence—which is perhaps best understood in terms of having a rational or autonomous will—while nonetheless having the kind of will that can play a nonderivative role in that person's well-being. My idea, simply, is that the threshold for being medically competent may well be higher than the threshold for whether one has the kind of will that can nonderivatively play a role in one's well-being. One might not have sufficient cognitive powers to count as a medical decision maker but enough of a picture of how one's life should go to have autonomy be a constituent part of one's well-being.

If that's right, then it is at least possible for a surrogate to consider the importance of following her charge's will in light of its direct or constitutive contribution to the charge's well-being and not just its non-

27. Brock and Buchanan identify self-determination as a stand-alone value and explicitly contrast it with the value of well-being (Buchanan and Brock, *Deciding for Others*, 29–40). I think this is a mistake: once self-determination is construed as a value, and not (just) as a right, it seems to make most sense to see it as part of the individual's well-being.

28. Of course, these do not necessarily mean the same thing, and there is much disagreement over what any one of these terms means. But the idea is just this: whatever having a rational or autonomous will amounts to, it is having this kind of will play a role in your life that is part of well-being.

derivative contribution. Indeed, it seems highly plausible to me that one thing we might aim for as a constituent part of the good of incompetent patients is that they shape their own lives to as great an extent as possible consistent with protecting or promoting their overall well-being.

Even if one rejects this idea, the first, uncontroversial point about how following an incompetent patient's will can contribute to his well-being is all I need for the central contention I'm making here, namely, that it can make perfect sense for the surrogate to give considerable weight—indeed enough weight to outweigh all other competing considerations—to Carl's will, to the point of doing what Carl wants entirely because he wants it (because it is good for him).

In such a case, Carl's will is not just substantially *relevant* but, as I shall put it, substantially *decisive* in determining the surrogate's decision.²⁹ The fact that Carl does not want surgery makes it the case that Carl's good lies in not having surgery because being able to enact one's will is a substantial good—and, we are imagining, decisively so—when weighed in the balance.

Importantly, this does not mean that we cannot identify Carl's good apart from what he wills. As I suggested above, it is perfectly intelligible to say, "Leaving aside Carl's will, getting the PEG is what is good for him." Now this may make it sound like Carl's good can be assessed apart from Carl's will so that when we factor Carl's will into the equation we are introducing some good-independent consideration. But that is not right: even though it is true that there is a sense in which we can assess Carl's good apart from his will, what we are doing here is simply making a partial assessment of Carl's good. It is partial precisely because it does not take into consideration the (possible) good of Carl's will being enacted. When we say, "Leaving aside Carl's will, getting the PEG is what is good for him," we are, in effect, saying, "Excluding consideration C, outcome O is best for Carl." The question then becomes: when we include consideration C, in this case the fact that he does not want to have the PEG, what is good for Carl?

In other words, an all-things-considered judgment of Carl's good will take into account the value of Carl's will being enacted. This means that if we assume that the focus of concern of Carl's surrogate is only Carl (and not also Carl's family, for example) then the only thing the surrogate can

29. It is, of course, possible that Carl's will figures as a weighty, or merely relevant, part of his good without being decisive in my sense. In such a case, we might call his will substantially relevant. Substantial decisiveness is the strongest form of substantial relevance. In what follows, I stick with the clearer case of Carl's will being substantially decisive, but my central point—about the difference between weighing someone's will as part of her good vs. considering it as a part of someone's right—goes through even if in almost all cases an incompetent person's will is merely substantially relevant. I thank an anonymous associate editor at *Ethics* for bringing this up.

properly care about is Carl's good. There can be no in-principle conflict between Carl's good and Carl's will, since Carl's will cannot ground practical reasons that have any normative force apart from considerations of his own good.³⁰ So, the well-intentioned surrogate cannot but act for Carl's own good no matter what she decides: there is no other kind of consideration that can properly be considered alongside Carl's good.³¹

We can now see the partial truth in each side of the dispute I mentioned at the outset about the proper role of an incompetent patient's wishes in determining treatment. The "no-weight" side is (presumably) responding to the fact that, unlike a competent patient's will, an incompetent patient's will does not ground authoritative demands. So his will really has no weight *qua* demand. But the "weight" side is right to respond that the patient's will still has some, perhaps decisive, weight as a part of the patient's good (either derivatively or nonderivatively). When we wonder, then, about the role of an incompetent patient's will in determining what to do, we need to be sure that we are not talking past one another by talking generally about the importance of "respecting autonomy." For in this dispute, one party is making the point that we cannot respect the patient's autonomy as right, while the other side is making the point that we can respect the patient's autonomy as part of his good. And these points are not only compatible but, I would suggest, correct.

III. THE DECISIVE REASON CRITERION

We have, then, a distinction between two ways of treating another's will as decisive: we can treat someone's will as either substantially decisive or as structurally decisive. This opens up the possibility that one can do what the other wants because he wants it, while nonetheless not treating his will as decisive in one sense. This is because one can treat the other's will as substantially decisive but not as structurally decisive. Such is the case with Carl and his surrogate.

This point helps us make good on the claim at the end of Section I, where I said that the real explanation for what makes the action of Bob's poorly behaved doctor, or of Ernest's benefactor Sabina, paternalistic is not that they act contrary to Bob's and Ernest's wills. Rather, it is that they do not

30. Strictly, they cannot have any force apart from considerations of the good, which need not include only his own good. For example, Carl might not want to have surgery because he does not want to burden his family any longer. This wrinkle does allow for Carl's will to conflict with his own good. But the central point is still intact: from the point of view of the surrogate, whose focus of concern is Carl, what Carl wills is only reason-giving for the surrogate insofar as it relates to Carl's good (regardless of Carl's motivation).

31. This is actually a little too strong, since Carl still has rights. But the idea is clear enough—nothing about what Carl wills in this situation grounds a good-independent consideration that the surrogate must take into account.

treat the will of Bob and Ernest as decisive in determining what to do. Our question now is this: which sense of “decisive” is at play in that explanation?

To answer that question, consider another variation on the case of Bob and his doctor. This time, Bob’s doctor does what Bob wants. And she does it because Bob wants it. But in this case, she treats Bob’s will as substantially decisive but not as structurally decisive. She considers Bob’s wish not to have the PEG as a component of Bob’s good and determines that Bob ought not to have the surgery. Suppose she tells Bob, “I have decided that you should not have surgery, because you do not want it.” We can imagine Bob being surprised, and upset, at this way of putting things: “What do you mean *you* have decided? It was not your decision to make!” The doctor here considered Bob’s will as part of a larger set of considerations about what is good for Bob when, in fact, Bob’s will should not have played that role in her deliberations. It should have acted as the decisive consideration in the sense of making questions about whether it is good for Bob to get the PEG irrelevant.

The modified Bob case, then, suggests that the notion of *structural* decisiveness is central to understanding the idea that Bob’s poorly behaved doctor (who either acts against Bob’s wish not to have a PEG or who treats his will as merely substantially decisive) and Sabina act paternalistically in virtue of not treating the wills of Bob and Ernest as structurally decisive:

DECISIVE REASON CRITERION: A acts paternalistically toward B when A acts for the sake of B’s good but does not treat B’s will as structurally decisive in determining what she, A, should do.³²

With the DECISIVE REASON CRITERION we can conclude that when Bob’s doctor treats Bob’s will as substantially decisive, but not as structurally decisive, she acts paternalistically. It is worth emphasizing once again that I am not interested in arguing over how we ought to use the term “paternalism.” Different authors mean slightly different things by it and use it for different purposes.³³ My central interest is highlighting how we can disrespect another person’s will, in acting for his good, even when we act in accord with his will because it is his will. And we have a clear case of

32. This conception of paternalism, like others, can take a more or less “internal” or “external” form. What I mean by this is that instead of putting things in terms of “acting for the sake of B’s good,” we might say instead, “acting for the sake of what A takes to be B’s good.” The same goes for “does not treat B’s will as structurally decisive.” One might say instead something like, “does not treat what A takes to be B’s will as structurally decisive.” Whether we should prefer an internalist or externalist conception of paternalism is orthogonal to my central concern.

33. Sometimes they mean very different things by it, as evidenced by Thaler and Sunstein’s use of the term, which is very expansive: “The paternalistic aspect [of libertarian paternalism] lies in the claim that it is legitimate for choice architects to try to influence people’s behavior in order to make their lives longer, healthier, and better” (R. H. Thaler

that here: in treating Bob's will as substantially decisive, but not as structurally decisive, Bob's doctor fails to respect Bob's will. She is not according it the place it ought to have in her deliberations.

We miss this point entirely if we work with a conception of paternalism that has the CONTRARY TO WILL CRITERION at its center. For whether Bob's doctor treats his will as substantially decisive or as structurally decisive, she does not act contrary to his will. In other words, the CONTRARY TO WILL CRITERION treats the case where Bob's will is treated as structurally decisive the same as the case where his will is treated as substantially decisive: in neither case did Bob's doctor act against his will. Therefore, neither acts paternalistically. Again, the term "paternalism" is not important here. The key point is that there is a morally important difference between the case where Bob's will is treated as substantially decisive and the case where it is treated as structurally decisive. That difference is easily missed if we appeal to the CONTRARY TO WILL CRITERION.³⁴

The DECISIVE REASON CRITERION then allows us to mark a morally significant difference that the CONTRARY TO WILL CRITERION does not. It brings into view, in a way that the CONTRARY TO WILL CRITERION does not, the way that someone can fail to respect the will of another person while acting in accordance with the other person's will because it is his will. Insofar as this is a phenomenon worth highlighting—and the importance of distinguishing between ways in which we can act in accordance with the will of another in medical decision-making contexts suggest that it is—we have reason to take the DECISIVE REASON CRITERION seriously.

It is also worth briefly noting that the DECISIVE REASON CRITERION is able to sidestep the two other problems we identified for the CONTRARY TO WILL CRITERION, namely, the No Will problem and the Accidental Concordance problem.³⁵ With respect to the No Will problem, the DECISIVE REASONS CRITERION allows us to count as paternalistic actions that are done for the good of another person but where that person has no views about the action (or kind of action), perhaps because he doesn't know (or hasn't thought) about it. In these kinds of cases, the person performing the action cannot treat the will of the other person as struc-

and C. R. Sunstein, *Nudge: Improving Decisions about Health, Wealth, and Happiness* [New Haven, CT: Yale University Press, 2008], 5). I'm inclined to think that this conception of paternalism is too far afield from mainstream use to warrant the name. The same, I think, is true of Savulescu's use of the term, in Julian Savulescu, "Rational Non-interventional Paternalism: Why Doctors Ought to Make Judgments of What Is Best for Their Patients," *Journal of Medical Ethics* 21 (1995): 327–31.

34. At this point, one might wonder whether the case of Carl causes problems for my view since, according to my conception of paternalism, the surrogate decision maker acts paternalistically in not treating Carl's will as structurally decisive but does not, thereby, disrespect Carl's will. I say more about this in just a little bit.

35. See above, Sec. I.

turally decisive, since the other does not have a will with respect to the matter at all.³⁶

With respect to the Accidental Concordance problem, the DECISIVE REASON CRITERION can easily declare paternalistic those actions where the paternalist's action accords with the paternalized subject's will accidentally (indeed, we were led to the DECISIVE REASON CRITERION by the problem of Accidental Concordance). For to treat another's will as structurally decisive in determining what to do is to settle on a course of action because it is what the other person wills. One cannot accidentally act (or intend to act) in a way that accords with the will of the other person if one is, in fact, treating that person's will as structurally decisive. The DECISIVE REASON CRITERION, then, avoids the problems that confront the CONTRARY TO WILL CRITERION while allowing us to mark a morally significant phenomenon—that of failing to respect someone by acting in accordance with his will because it is his will—that the CONTRARY TO WILL CRITERION does not.

But now we might have a worry. My central complaint against the CONTRARY TO WILL CRITERION is that it covers over a morally important difference between the case where Bob's will is treated as structurally decisive by his doctor and a case where it is treated as only substantially decisive. The DECISIVE REASON CRITERION, on the other hand, allows us to register this difference.

But where is Carl in all this? Carl, recall, is incompetent. So it would be inappropriate to treat his will as structurally decisive. Even so, Carl's surrogate decision maker may well treat his will as substantially decisive. According to the DECISIVE REASON CRITERION, she thereby acts paternalistically toward Carl. Indeed, she also acts paternalistically if she does not treat Carl's will as substantially decisive, so long as her actions are for Carl's good. This might strike some as odd, but I do not think it is. The very idea of making a decision where the only consideration is the other person's good is, I think, inherently paternalistic. This means that the role of surrogate decision maker for someone who lacks decisional capacity and does not have an advance directive of some sort is inherently paternalistic.

This would seem to raise the following worry for my view. Just as I claimed that appeal to the CONTRARY TO WILL CRITERION obscures a morally important difference, so too does appeal to the DECISIVE REASON CRITERION. For if failing to treat the will of another as structurally decisive for that person's own good is enough to render an action paternalistic, then, as just noted, both the doctor who treats Bob's will as substantially

36. One might worry that this response to the No Will problem threatens to make far too many kinds of actions paternalistic (e.g., ordinary cases of gift giving). I discuss this below, n. 37.

decisive and Carl's surrogate act paternalistically. Perhaps that is right—again, the label doesn't matter. But deploying the DECISIVE REASON CRITERION here covers over the morally important difference between the two cases—namely, that the doctor fails to respect Bob's will but the surrogate does not fail to respect Carl's will. Carl does not have the kind of will that ought to be treated as structurally decisive, and so the surrogate in no way disrespects Carl's will by treating it as merely substantially decisive. Bob's doctor's paternalism is (other things being equal) impermissible, while Carl's surrogate's is not. The DECISIVE REASON CRITERION cannot, on its own, account for this difference. Indeed, it seems to obscure it by labeling both actions “paternalistic.”³⁷

Let us be clear about what the problem is. One might think that an adequate conception of paternalism must be essentially normative, that it must account for the thought that paternalism is presumptively impermissible. If that's right, then assuming that there is agreement on the other elements of a conception of paternalism and that they don't bring the required normativity with them, the DECISIVE REASON CRITERION conception of paternalism has a problem. But so too does the CONTRARY TO WILL CRITERION. Neither is essentially normative.

It is hardly clear, however, that an adequate conception of paternalism must be essentially normative. Instead, and more plausibly in my view, we might think that an adequate conception of paternalism need not be essentially normative but that its content must make it reasonably

37. One might also have the additional worry that my conception of paternalism makes ordinary cases of gift giving paternalistic. Why? When I buy a friend a gift for no particular reason (to take the simplest case), I do so for my friend's good, and I do not take my friend's will as structurally decisive in deciding to buy a gift. Of course, I might buy the gift because I know my friend wants it (perhaps even because he wants me to buy it for him), but I do not see my friend's will here (or what his will would be) as grounding an authoritative demand. At best, I treat my friend's will as substantially decisive. If that's right, then according to the DECISIVE REASON CRITERION, I act paternalistically.

But I think that is not right. It is plausible to think that our ordinary practice of gift giving involves the giver taking herself to have and need tacit permission to buy gifts for the recipient, permission that is (implicitly) withdrawn if, for example, I insist on buying gifts that I think are good for my friend but that she hates or resents. If that is so, then I do treat my friend's will as structurally decisive.

Note, however, that if one rejects this move, the fact that gift giving ends up as paternalistic on my view is not damning since my central aim is not to capture ordinary intuitions about paternalism. “Paternalism,” recall, is acting in this essay as a placeholder for not treating another's will as structurally decisive and acting for the other's good. If, however, ordinary gift giving ends up being paternalistic on my account, then the issue I'm raising here—that only some instances of not treating another's will as structurally decisive count as failing to respect the other person's will—is all the more pressing, since there is certainly an important moral difference between me giving a gift to my friend and Bob's doctor treating Bob's will as only substantially decisive! If one is not convinced by the idea that the ordinary giver implicitly treats the recipients will as structurally decisive, then I think the account I give below (Sec. IV) deals with gift-giving cases. In fact, I think it explains all instances of permissible unasked for, yet helpful, behavior (of which gift giving, when done well, is an example).

clear why paternalism is presumptively impermissible.³⁸ With this in mind, one might hear the worry above as getting at the more general complaint that a conception of paternalism that involves the DECISIVE REASON CRITERION fails to make reasonably clear why paternalism is presumptively impermissible.³⁹

If that's the worry, then it is misplaced. Surely it is reasonable to think that once we understand what it means to treat the will of another person as structurally decisive, we will see that the onus is on the paternalist to justify her failure to treat the will of another as structurally decisive for that person's own good. For, presumptively, failing to treat the will of another as structurally decisive when it comes to that person's good is a failure to respect that person's will. In some cases—say, where a nurse practitioner is dealing with a patient in a psychiatric institution—that presumption is set aside very easily and more or less immediately (we might, in this case, say that the presumption is set aside at an institutional level and not on a case-by-case basis), but it is still there. So a conception of paternalism that includes the DECISIVE REASON CRITERION meets the demand that an adequate conception of paternalism must make it reasonably clear why paternalism is presumptively impermissible.

Insofar as the DECISIVE REASON CRITERION meets that demand, it also explains precisely why the doctor who treats Bob's will as substantially but not structurally decisive acts, other things being equal, impermissibly. Absent some account of why the impermissibility of acting paternalistically in this case is defeated, the DECISIVE REASON CRITERION makes clear why the doctor fails to respect Bob's will by acting in accordance with his will because he wills it.

But we're still left with the case of Carl, which is clearly one of permissible paternalism. And so we can hear the original worry as simply asking for an explanation of what explains the differing normative judgments in the case of Bob's doctor and Carl's surrogate. What makes it the case that one act of paternalism counts as a failure of respect for the other person's (Bob's) will, while the other does not?

IV. THE AUTHORITY AND THE SOUND JUDGMENT ACCOUNTS

The obvious place to look for a general account of what makes it the case that in some cases acting paternalistically is impermissible while in others

38. G. Dworkin offers a brief discussion of these two approaches to defining paternalism. Dworkin claims that, "As a matter of methodology it is preferable to see if some concept can be defined in non-normative terms and only if that fails to capture the relevant phenomena to accept a normative definition." I'm not sure that I accept this as a general methodological principle, but I think Dworkin is right in this case. See Dworkin, "Paternalism."

39. Of course, if one rejects this constraint on an adequate conception of paternalism, then there is no (even apparent) problem for my view.

it is not in the difference between Bob and Carl: Bob's will, but not Carl's, grounds an authoritative demand. So, we might say the following:

AUTHORITY account: Acting paternalistically is unjustified when the paternalizing agent does not treat another person's will as structurally decisive in determining what to do, even though the other person has a will that grounds an authoritative demand with respect to the issue at hand.⁴⁰

If we go this route, then Bob's doctor's paternalism is unjustified, but Carl's surrogate's paternalism is not. This is because Bob's, but not Carl's, will grounds an authoritative demand. So, Bob's doctor disrespects Bob's will in treating his will as merely substantially decisive, but Carl's surrogate does not disrespect Carl's will in treating his will as merely substantially decisive.

I think the **AUTHORITY** account deals with the case of Bob and Carl very well. It seems natural to say that Bob has the right to be the final decision maker when it comes to his health-care decisions, while Carl does not. More generally, in any case where we are inclined to say that someone has the right to make a decision, failure to respect exercise of that right for the person's own good (which can include treating the person's will as substantially but not structurally decisive) will, other things being equal, count as an instance of impermissible paternalism.

The **AUTHORITY** account does have merit. It explains the impermissibility of paternalism in cases where the paternalized subject, but not the paternalist, has a right to be the decision maker. But are all such cases of impermissible paternalism rights-involving in this way? Consider the following case:

ELEANOR'S NO-TRANS-FATS POLICY: Eleanor owns a large company. She is considering implementing a no-trans-fats policy in the cafeteria at work. She is motivated solely by her concern for the good of her employees. Before making her decision, Eleanor sends out a survey to all employees asking them to weigh in on the decision. The employees are unanimously opposed to the policy, although they don't have any particularly good reasons. Eleanor decides to enact the policy.

The **DECISIVE REASON CRITERION** tells us that Eleanor acts paternalistically in this case (as does the **CONTRARY TO WILL CRITERION**). Moreover, I

40. Gert and Culver are advocates of this kind of strategy, although they prefer talk of "moral rules" to "moral rights" on the grounds that the former is "clearer" (Bernard Gert and Charles M. Culver, "Paternalistic Behavior," *Philosophy and Public Affairs* 6 [1976]: 45–57, 52 n. 6). Feinberg's distinction between hard and soft paternalism suggests that he is also an advocate of the **AUTHORITY** account (Feinberg, *Harm to Self*, 12–16). Shiffrin is also an advocate of this strategy, although she builds the **AUTHORITY** account into the very definition of paternalism (Shiffrin, "Paternalism, Unconscionability Doctrine, and Accommodation").

assume that many will find Eleanor's action unjustified. If that's right, then we need to be clear about why this is. If Eleanor's decision, even if made after consulting her employees, was determined by concern for her bottom line, or in an attempt to decrease health-care costs for her business, or, as a matter of conscience, to extricate her company from contributing to what she sees as unacceptably unhealthy eating practices,⁴¹ then I suspect many of us, myself included, would say that Eleanor does not act impermissibly in implementing the policy. Indeed, if any of these are her reasons, Eleanor would not be acting paternalistically at all (according to both the DECISIVE REASON CRITERION and the CONTRARY TO WILL CRITERION), since she is not acting for the good of her employees.⁴²

To the extent, then, that we are right to assume that Eleanor acts impermissibly here,⁴³ it is because she is motivated solely by concern for her employees' well-being. She is taking advantage of her position of control over part of her employees' lives to make (or incentivize⁴⁴) them do something for their own good.

How can the AUTHORITY account explain this example of impermissible paternalism on the part of Eleanor? It would seem to have to say that the employees' (collective) will in this case grounds an authoritative demand. Of course, it might not ground an authoritative legal demand. But it grounds an authoritative moral demand. In failing to heed that demand in acting for their good, Eleanor fails to treat as structurally decisive a (collective) will that can, and should, properly be so treated.

Now this might strike some as strange: is it really so clear that the employees have the authority to demand that Eleanor not enact a policy at her company for their good? The advocate of the AUTHORITY account might well say "yes" and offer the following explanation in a Millian spirit: people's wills are, always and everywhere, structurally decisive with regard

41. This last case is, perhaps, less clear. Shiffrin ("Paternalism, Unconscionability Doctrine, and Accommodation") convincingly (in my view) argues that cases of conscientiously objecting to a practice should not be confused with acting paternalistically, let alone impermissibly.

42. What about mixed cases? Is it enough to make an action paternalistic that concern for the employees' good played some role in determining Eleanor's action? We would need to know more about the nature of the mix. If consideration of the employees' good was sufficient motivation but not, in fact, the entire motivation, then I am inclined to think Eleanor acts paternalistically. If it was necessary but not sufficient, I am less inclined to say Eleanor acts paternalistically. The same is true if it was neither necessary nor sufficient but weighed as an additional factor in favor of the policy. We might, in these kinds of cases, want to make more precise judgments than declaring the action paternalistic or not *simpliciter*. We might say that to the degree it was motivated by concern for the employees' good the action was, to that degree, paternalistic, or paternalistic in that regard.

43. If you think we are not right, just hold on: I consider that intuition below.

44. They could, presumably, bring their own lunch or go out.

to matters concerning their own good. This is obviously wrong if it means that whenever I will something for my own good then I have a legitimate claim to it—having \$100 from your wallet might be good for me, but my willing to have that money does not ground a legitimate claim to it! Likewise, Eleanor’s employees, it seems to me, don’t have a right to not have the policy implemented if it is motivated by a concern to reduce costs.⁴⁵

But there is a more plausible way to hear the idea that people’s wills are structurally decisive with regard to their own good. The idea is that you can legitimately demand that I not make decisions that are motivated solely by concern for your good. To say that I act solely out of concern for your good is just another way of saying that I act out of concern for your good without treating your will as structurally decisive (since, as we have already seen, my reasons for treating your will as structurally decisive do not have do with your good; they have to do with your right). Inasmuch as Eleanor acts solely out of concern for the good of her employees, she acts against their authoritative demand that she not implement the policy. So the *AUTHORITY* account explains why Eleanor’s action is impermissibly paternalistic.

The *AUTHORITY* strategist’s general account of when and why someone’s will is structurally decisive also allows us to explain the fact that while Bob might have the right to refuse surgery, he does not necessarily have the right to demand it. The doctor may have legitimate reasons for not doing the surgery that have nothing to do with Bob. What Bob can legitimately demand, on this account, is that the doctor not refuse to do the surgery simply from consideration of Bob’s good (given his will). Since Bob’s will grounds an authoritative demand, the doctor must take into account more than Bob’s good.

Our discussion of Carl, however, shows that we need to be careful if we go with the *AUTHORITY* account. For Carl’s will and, more generally, the wills of medically incompetent patients are not properly structurally decisive with respect to their own good. So it isn’t always true that people’s wills are structurally decisive with regard to their own good. The advocate of the *AUTHORITY* account, then, needs to say something like the following: people’s wills are structurally decisive with regard to matters concerning their own good when they have an intact or properly functioning will.⁴⁶

45. Of course, there might be other cost-cutting policies (refusing to pay for health insurance, for example) that they do have a right not to have implemented.

46. We might wonder what precisely is involved in having a “properly functioning” or “intact” will. One problem here is that having a diminished or defective will cannot simply consist of making poor choices, since we would then lose the distinction between Carl and Bob. Now one way to count as medically incompetent, or nondecisional, is to lack a will altogether—coma patients and very young children (basically, babies) may be said to lack

So, for the sake of clarity, the AUTHORITY account should be reformulated as follows:

AUTHORITY* account: Acting paternalistically is unjustified when the paternalizing agent does not treat the will of another, whose will is intact (and so grounds authoritative demands with respect to the person's own good), as structurally decisive in determining what to do.

Thus amended, the AUTHORITY* account gives us one way of capturing the normative difference between the Carl case and the case where the doctor treats Bob's will as substantially, but not structurally, decisive. More generally, it offers an account of why failing to treat the will of another person as structurally decisive, for that person's own good, is presumptively impermissible: people with intact wills have the authority to make decisions concerning their own good. Finally, it explains when the charge of impermissibility is defeated—namely, in those cases where we are dealing with people who lack intact wills (whatever that ultimately means).

Even so, I think the AUTHORITY* account faces two challenges. We can get at the first challenge by noticing that the AUTHORITY* strategy requires an extremely robust conception of the rights people have simply in virtue of having intact wills. The idea is that those with an intact will are, as Rawls puts it, “self-originating sources of valid claims”;⁴⁷ simply in virtue of having an intact will, my will is structurally decisive with regard to choices that have only to do with my own good.

Now this may seem unproblematic, indeed rather attractive, to some.⁴⁸ The challenge is just that it is not at all clear that our common-sense beliefs about which paternalistic actions are (im)permissible commit us to such a substantial view. It seems that someone could remain relatively agnostic about which rights people have while having relatively

wills altogether (although the latter certainly have preferences). Moreover, certain forms of mental illness and cognitive impairment are so severe as to make their sufferers lack a will (but not mere preferences) altogether. But it is too strong to lump all incompetent patients into these groups. It is possible for someone to have a compromised or defective will while still having a will on what we might call an “intuitive” conception of willing—the person can articulate a fairly consistent, coherent picture of what he wants and why he wants it. I imagine Carl to be such a person, but the case is too underdescribed to make this class of patient vivid. A good example of such a patient may be “Mrs. G.,” who is discussed at length in Jodi Halpern, *From Detached Concern to Empathy: Humanizing Medical Practice* (Oxford: Oxford University Press, 2001). In any case, the AUTHORITY account needs to offer an account of what makes a will intact since, without it, the strategy is highly implausible. But, as we will see below, my preferred strategy faces the same issue.

47. John Rawls, “Kantian Constructivism in Moral Theory,” *Journal of Philosophy* 77 (1980): 515–72, 543.

48. Contemporary versions of the idea, as I noted above, are defended by Feinberg (*Harm to Self*) and Darwall (*The Second-Person Standpoint*).

clear intuitions about whether a paternalistic action is justified. This is particularly apparent, I think, if we focus on the case of Eleanor and her employees. Rights talk seems out of place there.

But suppose it doesn't strike you as out of place. There is still the second challenge to the AUTHORITY* account: there are cases where treating someone's will as substantially decisive but not as structurally decisive for that person's good—acting paternalistically, as I am putting it—is intuitively permissible (i.e., is not disrespectful) even when the person has a perfectly intact will. For consider the following variant on the Eleanor case:

ZELLA'S NO-TRANS-FATS POLICY: Zella owns a large company. Zella's employees petition her to enact a no-trans-fats policy on the grounds that it will be good for them. Zella is persuaded that it would be good for the employees to have such a policy.⁴⁹ She enacts the policy.

Let us stipulate that Zella takes her employees' wills to be substantially decisive but not structurally decisive, although she admits that their wills are perfectly intact. Moreover, while she takes herself to have standing permission to enact the policies that her employees want enacted for their own good, she does not believe that she *needs* their permission to enact such policies (if she did, then she would take their wills to be structurally decisive with respect to their own good).⁵⁰

Notice that the DECISIVE REASON CRITERION tells us that Zella, like Eleanor, acts paternalistically. Unlike with the Eleanor case, however, I assume that everyone agrees that Zella is justified in enacting the no-trans-fats policy. At the very least, the Eleanor case is difficult in a way that the Zella case is not. But the Zella case is importantly different from the Carl case, since the paternalized subjects (the employees) have intact wills (and are taken to have intact wills by Zella). We have, then, a case where someone treats the intact will of another as substantially, but not structurally, decisive (for the other person's good) but where we are not inclined to say that the first person acted impermissibly (at least I am not inclined to say that).

One thing to say here is that while Zella's employees' (collective) will may be intact, it is not, thereby, authoritative. So, Zella is right to treat their will as substantially, but not as structurally, decisive. If that's right,

49. She might think it is good for them apart from their willing it. And she might think it is good for them in part because they will it, either derivatively (perhaps because enacting the policy will make her workers feel empowered, and that is good for them) or nonderivatively (perhaps because living their lives in accordance with their values is good for them). In any case, so long as she attending exclusively to the employees' good in making the decision she is, on my way of putting it, acting paternalistically.

50. This is what separates this case from the standard case of gift buying. See above, n. 37.

then Zella's paternalism is not impermissible according to the AUTHORITY* account, or at least according to what would have to be a modified version of it since, as it is stated above, the AUTHORITY* account yokes together having an authoritative will with having an intact will. But even leaving that problem aside, we're still left with another: denying authoritative status to the wills of Zella's employees deprives us of the judgment that Eleanor acts impermissibly. The AUTHORITY* account explained the latter judgment precisely by appealing to the idea that, in virtue of having intact wills, Eleanor's employees are authoritative with respect to their own good. But if that is true of Eleanor's employees, it is equally true of Zella's employees. So if Eleanor goes wrong in acting paternalistically, then so too does Zella. But the suggestion above is that Zella does not go wrong in acting paternalistically. Likewise, if we insist that Zella's employees' (collective) will is not properly authoritative in the above situation, and so Zella's paternalism is permissible, the same would need to be true of Eleanor's paternalism. But the suggestion above is that Eleanor does go wrong in acting paternalistically.

One response to this dilemma for the AUTHORITY* account would be to modify one of our judgments about the (im)permissibility of Eleanor's or Zella's actions. Zella's action seems the more amenable to reevaluation. Some might want to say that while enacting the policy is permissible, the manner in which Zella takes up the employees' request—as a kind of benevolent overseer of her employees' health at work (for their own good)—is impermissible for the kinds of reasons appealed to above in explaining the impermissibility of Eleanor's action. If that is right, then the two cases are on par, and the AUTHORITY* account explains the proper judgment in each case.

I, however, want to accept the initial judgments about the cases. If we do that, then the AUTHORITY* account can't be the right story, since there are some instances of (im)permissible paternalism that it can't account for. Or, more weakly, it cannot be the whole story.⁵¹ The account I favor doesn't pin the issue of permissibility to whether the paternalized subject has authority with respect to the choice/action in question. Instead, it pins it to the attitude of the person who does not treat the other person's will as structurally decisive.

Consider the respective attitudes toward their employees expressed by Eleanor and Zella. Eleanor takes herself to be acting for the good of her employees in light of what she sees as a deficiency (and likely

51. I am unsure myself whether the story I'm about to tell should be seen as an additional or alternative account to the AUTHORITY* account. I think it can function as an alternative, although I also think that it offers a less intuitive account of the impermissibility of paternalism when the paternalized subject clearly has a right to make the choice in question.

continuing deficiency) in their judgment. Zella too is acting for the good of her employees. But she is not acting because she takes their judgment in this matter to be unsound (indeed, she presumably thinks it very likely that their judgment is sound). This suggests the following impermissibility condition to accompany the DECISIVE REASON CRITERION:

SOUND JUDGMENT account: Acting paternalistically is impermissible when the paternalizing agent does not treat the will of another, whose will is intact, as structurally decisive in determining what to do, in part because she believes the other person is likely to fail to exercise a capacity for sound judgment in the situation at hand.

According to this strategy, Eleanor's paternalism is impermissible because it is motivated by a belief that her employees are likely to fail to exercise sound judgment on the matter of trans fats in the workplace. Zella's, on the other hand, is permissible precisely because she does not have this belief. On the contrary, she has every reason to think her employees would exercise sound judgment on the issue because they do!

The key question, then, for judging whether an instance of paternalism is (im)permissible turns on whether the paternalist is (partly) moved by the belief that her judgment is needed as a check or possible corrective on the judgment of the paternalized subject and that this is at least part of why the paternalized subject's will is not treated as structurally decisive. Talk of a "check" or "corrective" suggests that in order for an instance of paternalism to be impermissible, the paternalist need not believe that the paternalized subject is failing to exercise a capacity for sound judgment in a grave way or altogether but only that he is failing to exercise the capacity as well as the paternalist. The unjustified paternalist may take herself to simply have better judgment either in general or in the particular case at hand, perhaps because she thinks she has appreciated something that the paternalized subject has not.

But what will the SOUND JUDGMENT account make of the Carl case? The reason Carl's will is not treated as structurally decisive is precisely because the medical community and his surrogate believe that someone else's judgment is needed as a check on Carl's. We are assuming, safely I hope, that the surrogate's paternalism is permissible. But the SOUND JUDGMENT strategy would seem to tell us that it is not. It appears to lump together the Carl case and the case where Bob's doctor treats Bob's will as substantially but not structurally decisive. And this is precisely what we've been trying to avoid.

Contrary to initial appearances, the SOUND JUDGMENT account does not treat the two cases as the same. For it is not enough, according to this strategy, that the paternalist believes that the other person simply *lacks* a capacity for sound judgment. It must be that she believes that the person *fails to exercise* her capacity for sound judgment. Talk of "failure" is

normative: the idea is that the paternalized subject has the capacity for sound judgment and ought to exercise it but is not, in fact, doing so. This does not describe Carl, who we are imagining does not have a properly functioning will in the first place. In other words, he lacks the capacity for sound judgment in the situation at hand altogether. So while it is true that he does not exercise a capacity for sound judgment in refusing the PEG, he is not thereby failing to do so.

Bob, on the other hand, does fail to exercise his capacity for sound judgment in refusing the PEG. He is a competent medical decision maker, after all. Consequently, according to the SOUND JUDGMENT account, his doctor acts unjustifiably paternalistically if she treats his will merely as substantially decisive. And that's true even if she does what he wants because he wants it. Her attitude toward his will, in light of her belief that he has a capacity for sound judgment, is the wrong one. She disrespects him in doing what he wants because he wants it.

The SOUND JUDGMENT account, then, explains why paternalism is permissible (when it is) and why it is impermissible (when it is), and its explanation reaches further than the AUTHORITY* account. Even so, the SOUND JUDGMENT account is in need of further development. In particular, it would need to deal with at least two issues, one ethical and one metaphysical.

The metaphysical issue has to do with what it means for someone to have a capacity but to fail to exercise it. I am here relying on the claim that this idea makes sense and that we are right to see Carl's non-exercise of sound judgment as different in kind from Bob's: Carl cannot exercise sound judgment, while Bob could but does not.⁵²

The ethical issue has to do with why paternalism is unjustified when the paternalist believes that the other person will fail to exercise a capacity for sound judgment. Of course, it's clear what the problem is when the paternalist is unjustified in believing this. But oftentimes, the paternalist is *justified* in believing this. And even here, most of us think that paternalism is unjustified. Why should this be? If someone is really going to make a poor choice with respect to his own well-being,⁵³ why wouldn't it be justified to paternalistically intervene just because he has the (here unexercised) capacity for sound judgment? This is a hard

52. My account of how the normative status of paternalistic acts depends on whether the paternalized subject has (or is believed to have) the capacity to act/judge well bears a resemblance to Susan Wolf's account of moral responsibility in Susan Wolf, *Freedom within Reason* (Oxford: Oxford University Press, 1993). Like my account, Wolf's depends on the idea that people can have capacities they fail to exercise (specifically, the capacity to respond to the true and the good). Wolf offers an account of how this is possible in chap. 5 of *Freedom within Reason*.

53. And let's imagine that his own well-being is the only salient consideration (i.e., he is not justifiably sacrificing his well-being for the sake of some worthy cause).

question. Answering it would involve giving an account of what is valuable about autonomy.

That is a project for another day. Fortunately, the central point of this essay and its implications for surrogate decision making in a medical context do not depend on embarking on it. For whatever the full account is of what makes some instances of paternalism permissible and others (most) impermissible, we are left with the central conclusion that, in acting for the good of another person, we can disrespect that person's will while not acting contrary to it, indeed, while acting in accordance with it because it is his will.