

**Further Reflections: Surrogate Decision-Making When
Significant Mental Capacities are Retained**

Jennifer Hawkins
Duke University

In a recent article in this journal, Mackenzie Graham makes an important contribution to the literature on decision-making for patients with disorders of consciousness.¹ There he argues, and I agree, that decisions for unresponsive patients who are known to retain some degree of covert awareness ought to focus on current interests, since such patients likely retain the kinds of mental capacities that in ordinary life command our current respect and attention. If he is right, then it is *not* appropriate to make decisions for such patients by appealing to the values they had in the past, either the values expressed in an advance directive or the values recalled by a surrogate. There are two things I wish to add to the discussion. My first point is somewhat critical, for although I agree with his general conclusion about *how*, ideally, such decisions should be approached, I remain skeptical about whether his conclusion offers decision-makers real practical help. The problem with these cases is that the evidence we have about the nature of the patient's current interests is minimal or nonexistent. However—and this is important—Graham's conclusion *will* be extremely relevant if in the future our ability to communicate with such patients improves, as I hope it will. This leads to my second point. Graham's conclusion illustrates a more general problem with our standard framework for decision-making for previously competent patients, a problem that has not been adequately recognized. So in what follows, I'll explain this problem and offer some brief thoughts on how we could modify the framework to avoid the difficulties.

Graham is focused on the special case of surrogate decision-making for patients with "cognitive-motor dissociation"—patients who from the outside show no signs of awareness (though they have sleep-wake cycles), but who have nonetheless been found via brain scan to possess some degree of "covert" or internal awareness.^{2,3,4,5} The brain scans suggest these patients are aware of much

going on around them, but are unable to control their body *in any way* and equally unable to communicate. That such patients exist is a relatively recent discovery, only a little over a decade old.⁶ Before that, such patients were assumed to be in a vegetative state, entirely unaware of anything. Of course, there are still some patients who really are in a vegetative state, unaware of anything. But we now face the difficult task of distinguishing these two groups. The most detailed bedside examination can't tell us for sure who retains some degree of consciousness. Only some form of brain scan can do that, typically MRI.

As Graham reminds us, it is estimated that approximately 15% of behaviorally unresponsive patients nonetheless have some degree of covert awareness.^{7,8} Although this is a small subgroup of all unresponsive patients, the range of mental capacities found to date in this subgroup is impressive. It includes language comprehension, the ability to focus attention, the ability to select a response, and a degree of working memory, as well as some elements of executive function, and some ability to interpret other people via "theory of mind."^{9,10,11} Graham argues that the presence of these relatively more sophisticated mental capacities makes it inappropriate for decision makers to simply rely on an advance directive or on knowledge of the patient's past treatment preferences. Since the patient may well have *current* preferences of a reasonably sophisticated sort, we should instead focus on these. One way to put it is that the mental capacities retained by these patients are (or at least may be) enough like those to which we ordinarily accord concurrent respect that we ought to respect them here as well.

It is worth distinguishing at this point between three groups of patients. First, there are those unresponsive patients who have had no more than a thorough bedside exam, but who do not have access to brain imaging technologies. In the future, hopefully, more of these patients will gain such access. But right now all we can say is that a very small percentage of these patients are likely aware to some degree, but we have no idea which patients these are. Second, there are patients who have been tested with brain scans and found to have covert awareness, but who are still not able to communicate with us in any way. This group will most likely grow with time as more patients gain access to the technology. Finally, there is a

tiny handful of patients to date who are known to be aware and who have successfully answered some 'yes' and 'no' questions via mental responses detected by brain scan.¹² However, we simply don't know if such patients will ever comprise more than a tiny percentage of those with covert awareness.

Thus Graham's conclusion about current best interests really only applies once a patient is relatively stable *and* is known to have some degree of covert awareness. In the immediate aftermath of injury no one knows whether a patient is aware or may eventually become aware. And for those stable patients with no access to anything but bedside exam, we can't tell if they are covertly aware. All we know is that most such patients are not. Graham's conclusion thus has direct application only in cases where some time has passed, a patient has stabilized, and we have evidence of covert awareness. Though this group is small at present, it will likely grow. The problem for his claim, however, is that even here, where covert awareness is known to exist, communication is usually not possible. Graham is right that patients with covert awareness *may* retain certain more sophisticated attitudes about their current situation, and right to think that *in principle* such attitudes would be relevant to treatment decisions if known. But currently they aren't known or indeed, *knowable*.

If we accept Graham's point that it is current attitudes that matter, the most we can do is hypothesize on general grounds about what the patient as she is now most likely wants. But whether or not this is a good policy depends on whether we can back up our hypotheses with good arguments *and* whether the best hypotheses all point towards the same kind of treatment decisions. It won't help us to have equally compelling arguments pointing to opposite conclusions about how to make choices for such patients. But I see no clear directions for treatment emerging from Graham's article.

Elsewhere, in an earlier paper¹³, I argued for a certain hypothesis about what the experience of many (even if not all) such patients is like. And my hypothesis had a clear implication for surrogate decision-makers. I hypothesized that many (even if not all) such patients probably suffer a great deal. In referring to suffering, I did not mean to imply that such patients are in physical pain (though of course that would

be an important thing to know and an important thing to address if it were true). Rather I meant that they likely experience great emotional distress and/or depression. My reasoning was not based on some general assumption that it must always be bad to be disabled or to be unable to do many things one did before. I am well aware that many able bodied people have such biases, and think that they would not want to live with particular disabilities. But I also know that people living with those same disabilities by and large report high quality of life, and I take that to heart.^{14,15} My reason for thinking things will be different in this one kind of case is that this disability is distinct from every other known disability: for no other disability, no matter how severe, brings such total isolation and powerlessness. And it is an isolation and powerlessness that cannot (at least currently) be relieved. As Graham notes (and as I noted in my earlier article), even patients with locked-in syndrome are able to communicate in some way, albeit often in a very laborious way, e.g. by blinking an eye. But so far the vast majority of patients with covert awareness are not able to communicate at all. And along with the loss of all bodily control, the complete inability to communicate would, I hypothesize, result in incredible feelings of powerlessness. Since there are known to be strong associations between isolation, feelings of powerlessness, and depression,^{16,17} I hypothesized that many (though not all) such patients would be depressed. And since we have no way to help them (if we did they would no longer be so isolated), I suggested that surrogates should take seriously the possibility that it might be best to let such patients die.

Graham thinks this is too strong, and worries that there will be much variation among patients in the ways they respond to their situation. Though I am not convinced, I am aware that many share his view. Fair enough. But then I ask: What if he is right? We would then need to adopt a more individualized approach. But we have no materials to help us hypothesize in a more individualized manner. So again, our lack of knowledge of what really goes on inside these patients makes decision-making incredibly difficult.

Despite current practical dilemmas, I do think Graham's argument is important. It is not bizarre to think that at some point we may gain the ability to

communicate to some degree with these patients. But if and when we do, we may well find ourselves in the following position: The kind and degree of communication possible is still not enough to enable us to determine whether the patient has decision-making capacity. So a surrogate will continue to be in charge. But because some communication has succeeded we now have more information from the patient about the patient's current experiences, and preferences. In such a case, Graham's claim that we ought to be giving most weight to current interests will tell us to pay attention to whatever we can learn from the patient. And that seems absolutely right.

This leads me to the second point I wish to make. As I see it, what is most significant about Graham's argument is that it points to a problem that comes up in other contexts as well, and which (so far as I know) has not received *general recognition as the problem it is*. The problem is that it is quite possible for an individual to lose decision-making capacity and yet retain many relatively sophisticated mental capacities, capacities that most of us value and see as central to who we are. Intuitively, when these capacities and attitudes remain in a patient, it seems they should guide our decision-making for the patient. Yet many theorists assume that past values and preferences of an individual (whether written in an advance directive or simply known to a surrogate) become authoritative as soon as the patient loses decision-making capacity.¹⁸

However, this idea can be challenged. To be clear, the challenge does not extend to the basic thought—accepted by many and enshrined in current law—that it is important to grant ordinary people some degree of precedent control—i.e. some power to make choices *now* for their own possible future. Nor does it challenge the idea that, at least sometimes, it is appropriate to make decisions for previously competent patients by appeal to their past values. One can accept all of this while acknowledging that there is a further, highly important question still to settle: namely, when or at what point do past values become authoritative for present decision-making? What capacity or set of capacities must an individual lose, before we shift to making decisions in this way? Many theorists assume that the only important capacities are those related to decision-making, and so once a patient

becomes globally incompetent (no longer able to make any of her own medical decisions), past values should direct choices.¹⁹ But this picture is too simplistic.

In his article, Graham refers to and discusses the work of Ronald Dworkin who has given one of the most influential arguments for the importance of advance directives. Dworkin focused his discussion on the case of Alzheimer's dementia, but he intended his conclusions about the normative force of advance directives to be general.²⁰ Like most theorists, he assumed that directives should be appealed to once a patient is globally incompetent.²¹ But even in the case of Alzheimer's dementia, there are good reasons to think Dworkin was wrong about this. Dworkin assumed that by the time an Alzheimer's patient is globally incompetent, she has lost any sense of herself as an individual whose life has a certain narrative structure.²² He also assumed that such an individual would have lost any memory of, or concern for, her past values, and would have lost the ability to develop new values.²³

In a powerful reply to Dworkin, Agnieszka Jaworska pointed out that patients can become globally incompetent long before they lose all of their old values. It might be true that they have lost the ability to form *new* values. And it might be true that such individuals no longer retain a sense of themselves as individuals whose lives have (or had) a certain narrative structure. But it is possible for individuals to lose that, and yet retain current values and concerns and desire to engage the world in ways dictated by those values.²⁴ When these capacities remain we should attend to their current concerns and *not* their past values.²⁵ She offered as an example the case of a woman who loved contributing to scientific progress and helping to improve care for other, future patients.²⁶ This woman was proud to be participating in a clinical trial for Alzheimer's, despite the fact that she got confused easily and had a hard time holding and manipulating information for very long (something one needs to be able to do to qualify as having decision-making capacity). In Jaworska's words, she was still very much a *valuer*.²⁷ And Jaworska argued that we should not appeal to advance directives when making decisions for such patients. Instead, directives should only become operative once the patient has ceased to be a valuer.

As I see it, Graham's argument is similar to Jaworska's, but focused on a different patient population. In both cases, concerns arise because certain patients who are deemed globally incompetent nonetheless have current attitudes and preferences that seem to demand our current respect. In both cases, because patients retain these current capacities, it seems inappropriate to base decisions for them on past values. It seems to me that Jaworska and Graham have each pointed to an issue that is probably much broader and, as yet, insufficiently explored. If we have these intuitions in these cases, then clearly we need to think more carefully about what capacities must be lost before past values become authoritative. But if we agree that this is important, how do we square this with the currently accepted rationale for honoring past preferences? For it is not random that theorists assume past values begin to matter as soon as global incompetence sets in. That assumption is built into the common rationale.

We can see why that is by looking again at a famous argument given by Dworkin. Dworkin explains the importance of following advance directives by emphasizing what he calls respect for autonomy. What Dworkin means when he talks about "respecting autonomy" is simply *honoring the free (i.e. non-coerced) choices of competent adults*.²⁸ Because it is morally important to respect the free choices of competent adults, he thought that once competence was lost, we continued to show the proper moral concern for the individual by honoring her past free choices—the free choices she made when she was competent. Here is how he argued for this.

He asked his readers to consider why we ordinarily take it to be morally important to honor the free choices of competent adults. The moral rationale for this is not (he argued) that competent individuals will make good choices for themselves. Many will make bad choices—bad even by their own lights.²⁹ The rationale for honoring the free choices of competent adults is different. According to Dworkin, it matters because it is only by giving people this power that we enable them to be the authors of their own lives, or to engage in "self-creation" (something Dworkin viewed as much more important than having one's life go well from a welfare standpoint).³⁰ If we accept that he is right about this rationale then, he

thought, we should see that the very same rationale supports giving competent adults the power *now* to make choices for *later*, and requiring others to honor those choices later. And since this authorial power is explicitly viewed as a power of competent adults, it makes sense that an advance directive would come into effect at the moment when this authorial power is lost, i.e. as soon as a patient becomes globally incompetent.

As I see it, everything turns on this vision Dworkin articulates. What matters to Dworkin is being *the author of one's own life*, or *creating a life with a certain narrative structure*. While I have no doubt that there are people who care deeply about these types of goals, I nonetheless think most competent adults have other less grandiose concerns as well. And it is these equally important concerns that ground their interest in being able to decide for themselves. So in order to make room for the kinds of insights unearthed by Jaworska (for patients with Alzheimer's dementia) and Graham (for patients with covert awareness), we need to reject Dworkin's oversimplified vision of why it is important for ordinary competent individuals to have a say over what happens to them, thereby changing his argument at the very first step.

In short, I think we should drop Dworkin's focus on self-creation, and on the long-term aspects of a good life, and acknowledge that there are many significant aspects of being an agent and of being a valuer that have nothing whatsoever to do with these more complex aims. We should ask ourselves the following question: what are the capacities we ordinarily possess that ground our deep personal interest in being decision-makers for ourselves? Presumably there are a number of these, and it may take time for theorists to come to agreement about all of them. But we should then recognize that some of these capacities can be lost even while others are retained. It is undeniable that when a patient loses decision-making capacity a surrogate must make choices for that patient. But this does not necessarily mean that the surrogate should appeal to past values or to an advance directive. Rather, we should say that the past only becomes relevant to present decision-making when all or most of these valued capacities are lost. In the meantime, if decision-making capacity is lost while the patient remains someone who values, or someone who

loves, or someone who clearly has current concern for what is happening around her, a surrogate should attend to what is known (or can be known) about these values and concerns and use *them* as the guiding basis for decision-making.

I have discussed Dworkin here as one highly prominent example of a theorist who defends the idea that past wishes become authoritative as soon as a person becomes globally incompetent. But of course other defenses are possible. Nonetheless, what most have in common, on my view, is an overly complicated picture of what it is that grounds our ordinary interests in being our own decision-makers. I don't know whether Graham would agree with what I have said here, but I do hope that in future theorists will start to explore more systematically the kinds of capacities that underlie this basic interest, so that we can move to a more nuanced view of how to decide for previously competent patients who are still "there" in some very important sense.

-
- ¹ Graham, M. Precedent autonomy and surrogate decision-making after severe brain injury. *Cambridge Quarterly of Healthcare Ethics* 2020; 29 (4): 511-26.
 - ² Monti M, Vanhaudenhuyse A, Coleman MR, Boly M, Pickard JD, Tshibanda L, et al. Willful modulation of brain activity in disorders of consciousness. *New England Journal of Medicine*. 2010; 362:579-89.
 - ³ Cruse D, Chennu S, Chatelle C, Bekinschtein TA, Fernandez-Espejo D, Pickard JD, et al. Bedside detection of awareness in the vegetative state: a cohort study. *The Lancet*. 2011;378(9809):2088-94.
 - ⁴ Kondziella D, Friberg CK, Frokjaer VG, Fabricius M, Moller K. Preserved consciousness in vegetative and minimally conscious states: systematic review and meta-analysis. *Journal of Neurology, Neurosurgery, and Psychiatry*. 2016; 87(5):485-92.
 - ⁵ Schiff ND. Cognitive motor dissociation following severe brain injuries. *JAMA Neurology*. 2015;72(12):1413-15.

-
- ⁶ Owen AM, Coleman MR, Boly M, Davis MH, Laureys S, Pickard JD. Detecting awareness in the vegetative state. *Science*. 2006;313(5792):1402.
- ⁷ See note 2, Monti et al.; 2010, at 579-89.
- ⁸ See note 4, Kondziella et al. 2016, at 485-92.
- ⁹ Fernandez-Espejo D, Owen AM. Detecting awareness after severe brain injury. *Nature Reviews Neuroscience*. 2013; 14(11): 801-09.
- ¹⁰ Naci L, Cusack R, Anello M, Owen AM. A common neural code for similar conscious experiences in different individuals. *Proceedings of the National Academy of Science USA*. 2014; 111(39): 14277-82.
- ¹¹ Graham M, Naci L, Owen AM, Weijer C. Covert narrative capacity: Mental life in patients thought to lack consciousness. *Annals of Clinical and Translational Neurology*. 2017; 4(1): 61-70.
- ¹² See note 9, Fernandez-Espejo, Owen 2013, at 801-09.
- ¹³ Hawkins J. What is good for them? Best interests and severe disorders of consciousness. In: Sinnott-Armstrong W, editor. *Finding Consciousness: The Neuroscience, Ethics, and Law of Severe Brain Damage*. Oxford: Oxford University Press; 2016. pp. 180-200.
- ¹⁴ Albrecht GL, Devlieger PJ. The disability paradox: high quality of life against the odds. *Social Science and Medicine* 1999; 48:977-88.
- ¹⁵ Ubel P, Loewenstein G, Jepson C. Whose quality of life? A commentary exploring the discrepancies between health state evaluations of patients and the general public. *Quality of Life Research* 2003; 12:599-607.
- ¹⁶ Teo AR, Choi H, Valenstein M. Social relationships and depression: ten-year follow-up from a nationally representative study. *PLoS One*. 2013;8(4):e62396. Published 2013 Apr 30. doi:10.1371/journal.pone.0062396
- ¹⁷ Rubenstein LM, Alloy LB, Abramson LY. Perceived control and depression: Forty years of research. In: Reich JW, Infurna FJ, eds. *Perceived Control: Theory, Research, and Practice in the First 50 Years*. New York: Oxford University Press; 2017; 229-52.
- ¹⁸ Hawkins J, Charland LC. Decision-making capacity. In: Zalta EN. ed. *Stanford*

Encyclopedia of Philosophy. URL = <<https://plato.stanford.edu/archives/fall2020/entries/decision-capacity/>>. Throughout, I treat “decision-making capacity” (or “capacity”) and “competence” as equivalent ways to talk about the same thing. Not all authors do so, but the common view that one term is clinical and the other legal is not really helpful for a variety of reasons. To give just one example, in the UK “capacity” is often said to be a legal term whereas the same is said about “competence” in the US.

- ¹⁹ It is important to distinguish between local and global incompetence/ incapacity. Most ethicists accept now that decision-making capacity should be assessed on a decision-by-decision basis, such that a patient might in principle be deemed incompetent to make one decision but found competent to make others. A patient like this has *local* incompetence. However, some patients are so seriously impaired that they are really unable to make any medical decisions for themselves. These patients are said to be *globally* incompetent. Although Dworkin is not nearly as clear about this as he should be, most of the time advance directives are assumed to come into play once a patient is globally incompetent. See Note 17, Hawkins, Charland 2020.
- ²⁰ Dworkin R. *Life's Dominion: An Argument about Abortion and Euthanasia*. New York: Viking; 1994.
- ²¹ Interestingly, Dworkin also defended a view of best interests according to which current interests of an incompetent patient are determined by past wishes. If one holds this view then, even if one grants that *interests* should guide us here (as opposed to respect for past competent decisions), this will not make a difference to how one proceeds. Past values of the patient will still be most important. I have argued against this view in Hawkins J. Well-being, time and dementia. *Ethics* 2014; 124: 507-42.
- ²² See note 20, Dworkin 1994: 230.
- ²³ See note 20, Dworkin 1994: 230-31.
- ²⁴ Jaworska A. Respecting the margins of agency: Alzheimer's patients and the capacity to value. *Philosophy and Public Affairs* 1999; 28 (2): 105-38.

²⁵ See note 24, Jaworska 1999: 135.

²⁶ See note 24, Jaworska 1999: 117-18.

²⁷ See note 24, Jaworska 1999: 130.

²⁸ See note 20, Dworkin 1994: 222.

²⁹ See note 20, Dworkin 1994: 223.

³⁰ See note 20, Dworkin 1994: 224.