

Jennifer Hawkins

Pre-print of Target Article American Journal of Bioethics, Published Online June 26, 2023

DOI: 10.1080/15265161.2023.2224273

Affect, Values and Problems Assessing Decision-Making Capacity

Jennifer Hawkins, PhD
Associate Research Professor
Department of Philosophy
Duke University

Address:
Department of Philosophy
Duke University
Campus Box 90743
Durham, NC 27708

Email:
jen.hawkins@duke.edu

Affect, Values and Problems Assessing Decision-Making Capacity

Abstract

The dominant approach to assessing decision-making capacity in medicine focuses on determining the extent to which individuals possess certain core cognitive abilities. Critics have argued that this model delivers the wrong verdict in certain cases where patient values that are the product of mental disorder or disordered affective states undermine decision-making without undermining cognition. I argue for a re-conceptualization of what it is to possess the capacity to make medical treatment decisions. It is, I argue, *the ability to track one's own personal interests at least as well as most people can*. Using this idea, I demonstrate that it is possible to craft a solution for the problem cases—one that neither alters existing criteria in dangerous ways (e.g. does not open the door to various kinds of abuse) nor violates the spirit of widely accepted ethical constraints on decision-making assessment.

Key Words:

decision-making, medicine, philosophy, psychiatry/psychology, professional-patient relationship

Article Text:

In contemporary western society, people generally value control over their lives. They value making their own medical decisions. Because of this, in contemporary western medicine patient authority over treatment decisions is a default assumption. Informed consent—an important part of what helped to make patient authority real—has been a foundational requirement in treatment and research for more than forty years. (Faden & Beauchamp 1986; Kim 2010, 6-11). And yet, sadly, not everyone is in the position to give informed consent or make meaningful decisions for themselves. The ethical practice of giving patients authority requires that we also have an ethical way of determining who can and who cannot meaningfully exercise such authority. The moral stakes in such determinations are very high.

The problem of how to ethically draw this line is referred to as the problem of how to determine whether someone has “decision-making capacity” (often just referred to as “capacity”) or

whether someone is mentally competent (often just shortened to “competent”).¹ If I am deemed to have the capacity to make a particular medical decision, then in most settings my decision will be honored regardless of what others think of it, and even if it strikes many as unwise. But if I am not deemed to have capacity I will not get to choose. Informed consent will be obtained from someone else with authority to decide *for me* (Buchanan & Brock 1990, 27). It thus matters a great deal how we *conceptualize* decision-making capacity and how we assess it. Not only is it a terrible thing to remove choice from someone who wants it and could exercise it appropriately, it is equally terrible to leave choice in the hands of someone who cannot exercise it appropriately. The first kind of mistake denies a person valuable freedom. The second abandons a vulnerable person to likely harm (Buchanan & Brock 1990, 40-41; Kim 2010, 3-4).

Given these stakes, it is surprising that there is nowhere near as much agreement about what decision-making capacity is or how to assess it as there is agreement about what informed consent is, and how to tell when valid consent has been obtained. There is one practical framework—a way of conceptualizing and assessing capacity—that has been developed with care and which has become dominant insofar as anything has (Kim 2010, 19-20). This is what has come to be known as “the four abilities model” developed by Paul Appelbaum, Thomas Grisso and colleagues (Appelbaum & Grisso 1988, 1995; Grisso & Appelbaum 1995, 1998a; Grisso, Appelbaum, Mulvey & Fletcher 1995). Nowadays, many clinicians in the U.S. and other places rely on it. No doubt many people assume that the prevalence of the model means that foundational questions about what capacity is, and how we ought to assess it, have been laid to rest. However, this is not the case.

There are certain kinds of cases (described below) that the framework is unable to handle.

¹ In the U.S., some think “capacity” refers to a clinical judgment whereas “competence” refers to a legal one. In the UK, insofar as a distinction is made, it is typically the opposite: “competence” is assumed to refer to a clinical judgment and “capacity” to a legal one. I use them interchangeably. I do not find distinguishing them helpful given that (1) even courts often use the terms interchangeably and (2) clinical judgments often have legal force (Kim 2010, 17-18) and (Hawkins and Charland 2020, §1.1).

However, to date no promising proposals have been made about how to change the framework to accommodate these. I believe this is because in order to solve the problem we must first re-consider the most foundational issues, namely, what precisely we are trying to determine when we assess capacity and what the ethical constraints on assessment are. In what follows, I first re-visit the kinds of cases that, intuitively, create problems for the existing framework. I then argue for one conceptual and one practical conclusion. I argue first that we should re-think what medical decision-making capacity really is. I defend the claim that, properly understood, medical decision-making capacity is *the ability to track one's own personal interests at least as well as most people can*. Second, I suggest a practical solution—an amendment to the existing framework that builds on this insight to solve the problem cases. Finally, I defend the proposal against various kinds of ethical worries. It is worth noting, however, that because the proposal is new, any attempt to actualize it would, of course, require time, further research, and efforts at refinement. Before anything like that can occur, people must see the promise in the approach. My current goal is not, therefore, to suggest that anyone immediately adopt my proposal. Without careful development of the ideas and collective agreement on interpretation of key elements, that would be ethically problematic. My goal is the more modest one of demonstrating the promise of this approach in the hope that others will agree that it is worthy of further inquiry, research and development.

§1.0 The Current Framework and Its Problems

Although different approaches to assessing capacity still exist, there is, as noted, one approach that has the best claim to being dominant. Much of what I say applies to other approaches as well, but I shall here simply focus on the four abilities model developed by Paul Appelbaum, Thomas Grisso and colleagues.² They have not only developed a framework listing four key abilities, but have also

² There remains disagreement about the extent to which the four-abilities model is consistent with legal requirements outside the U.S., in particular the UK Mental Capacity Act of 2005. However, it is clear that the

developed guidance for clinicians and researchers on how to interpret and assess these abilities (Grisso and Appelbaum 1998a), and empirical instruments to help assess the degree to which a person has each of the four abilities. The primary instrument is the MacCAT-T, which stands for MacArthur Competence Assessment Tool—Treatment (Grisso, Appelbaum & Hill-Fotouhi 1997; Grisso & Appelbaum 1998b). There is also a parallel instrument for research, the MacCAT-CR (Appelbaum & Grisso 2001). The model has been widely adopted for a variety of reasons. It was deliberately designed to fit well with current law (Berg et al. 1996). It has also been extremely well worked out, right down to very concrete details of assessment. Finally, the existence of an instrument means that it is relatively easy for clinicians to learn and use and fairly easy to obtain cross-context consistency.

However, a recurring criticism of the model is that it fails to adequately account for the ways in which a person's emotions and/or values can shape decision-making in problematic ways (Bursztajn et al. 1991; White 1994; Elliot 1997; Charland 1998a, 1998b, 2006; Halpern 2001, 2011, 2012; Breden & Vollmann 2004; Tan, Stewart, Fitzpatrick & Hope 2006; Vollman 2006; Hermann 2016). In general, the critics have not doubted that the four abilities are necessary for competence. Nor do I. However, I am convinced, as are many of the critics, that the four abilities are not always *sufficient*. There are a number of patients who, intuitively, should *not* count as having decision-making capacity, but who nonetheless possess the relevant cognitive abilities to a high enough degree that they count as having capacity on the four-abilities model.

The four abilities themselves are (1) the ability to evidence a choice, (2) the ability to understand, (3) the ability to appreciate, and (4) the ability to reason. I shall explain each briefly.

The first ability—the ability to evidence a choice—is the least mental of the four. It is mainly included in the model to remind clinicians that no matter what other capacities a patient may have, a patient must be able to come to some decision or other and clearly communicate it if others are to

two are close, and that both emphasize cognition in a similar way. Thus the concerns raised here plausibly apply to it as well.

honor it (Grisso & Appelbaum 1998a 34-37). Since this ability is not relevant here, I shall not mention it again.

The second ability—the ability to understand—requires that the patient be able to grasp all the facts relevant to her decision (Grisso & Appelbaum 1998a 37-42). This is usually tested by talking to a patient about the decision she faces, giving her information and asking her to repeat it back in her own words to ensure that she has not just memorized what was said to her. Follow up questions will likely probe her ability to draw out obvious implications of what she has learned.

The third ability—the ability to appreciate— requires, in addition to grasp of information, that a person also believe that the information is true of her (Grisso & Appelbaum 1998a, 42-52). This may not be what lay people assume “appreciation” means, but it is how it is usually interpreted in capacity assessments. Having such a requirement is important because cases do arise in which patients are able to grasp *what* is being said but refuse to believe that it applies to them. An example would be a patient with ICU psychosis who grasps that his doctors are telling him he is seriously ill but who believes he is just fine and they are not really doctors. If a patient fails to believe that the medical facts apply to him and if (as in this case) his failure is based on *delusion* as opposed to, for example, religious beliefs (Grisso & Appelbaum 1998a, 47-48), then he lacks decision-making capacity.

Finally, the patient must be able to reason to some degree. This is generally interpreted in a very minimal way such that many people—certainly many philosophers—would hardly recognize it as a reasoning requirement. It includes a number of related abilities, but primarily the ability to consider several possible outcomes of a decision, and some ability to see what these imply for one’s own values and concerns (Grisso & Appelbaum 1998a, 52-58).

To see why possessing these capacities is not always sufficient, consider the following two cases. First, consider a young man I will call Terence who was diagnosed two years ago with

anorexia nervosa.³ He was stable for a while, but he has now started losing weight again, bringing him to extremely dangerous weight levels. He is likely to die if he does not allow himself to be medically fed to gain weight. But he refuses, saying that he knows he has an illness, he knows he is incredibly thin, and he knows that he may die. But, he insists, he would rather die than put on weight. Does he have the capacity to make such a decision right now? According to the MacCAT-T and personal interviews, he does. He is therefore not hospitalized, and he dies several weeks later in his apartment of complications from starvation.⁴

Terence counts as having capacity *because* he understands the basic facts about his situation and accepts that this information applies to him. It is significant that he says he knows he is ill, for this means he has what psychiatrists call “insight.” He can reason well enough to grasp the different paths open to him, and what the results of each path might be. He knows there is a high chance of death without treatment. Finally, given his stated preference for avoiding weight gain, even at the risk of death, his choice is, broadly speaking, instrumentally rational—it makes sense given the values he currently has and the priority he gives them.

Now consider a woman in her mid-50s I shall call Donna—a modified version of a case highlighted in the work of Jodi Halpern (2001; 2011; 2012). Donna has type-1 diabetes that has over time increasingly given her problems. Yet despite her illness, she enjoys her life. A few years ago she had to have an above-the-knee leg amputation. She was initially depressed, but responded very well to treatment and eventually adjusted fully to her new situation. She has an interesting career as a visual artist, which is not affected by her disability. She is also socially active, with a number of close friends. She now learns that she will need to have a second above-the-knee leg

³ Although anorexia nervosa is more common among females, males make up approximately 10% of patients (Weltzen 2016).

⁴ The case of Terence is based loosely on cases described by Tan et. al. (2006a) in a small study of capacity among anorexia patients. All were deemed competent on the basis of assessment with the MacCAT-T. Although two patients clearly lacked appreciation, the rest seemed to have insight into their illness but, like Terence, were simply not willing to gain weight. One of the patients in this study says that death is preferable to gaining weight (2006a, 274-5).

amputation. Unlike before, she initially responds well to this news. She seems to accept it and plan for it. She comes in willingly for her surgery, with what seems like a good outlook. However, immediately afterwards she informs her care team that she wants no further treatment, except comfort care. She is clear that this includes all standard post-op medications including blood-thinners and routine antibiotics for the prevention of post-surgical infection. She insists that if she develops an infection she wants to let it run its course, even if she turns septic and dies. Puzzled by her change in attitude and the forcefulness of her demands her doctor calls for a psychiatric consult. The resident who comes to interview Donna discovers, after some probing, that Donna's husband visited her in hospital and told her he was leaving her for someone else and that he would be moving out of their home while she was in hospital. Obviously, this sheds new light on her change of outlook and her refusals. She is presumably reacting to this news with grief and despair. In lay terms we would say she is not *just* grieving, but is "in shock." Because her emotions are in turmoil, she is not currently attending to, or giving weight to, her future as it would most likely be. The question, however, is whether she has the capacity *right now* to make such a consequential decision. She is assessed by the resident according to the four abilities model and found to have capacity. She develops a serious infection soon after, and is allowed to die in hospital several days after that.⁵

Like Terrence, Donna is deemed to have capacity because despite emotional turmoil and its effects on her thinking about the future, her basic cognitive capacities are unaffected. She understands her current options and what will follow if she accepts care and if she declines. She understands that these options really apply to her, and she can relate them to her current concerns, which in this case are dominated by her present desire not to feel as abandoned as she currently does.

⁵ Donna is an altered version of "Ms. G" presented in Halpern (2001, chap 1). I deliberately changed the original to remove elements that, in my experience, tend to distract people from the issue of decision-making capacity.

Although there are various ways to explain precisely what is happening with Donna, I find plausible the description offered by Halpern (2012), who argues that in such circumstances a person develops a “concretized emotion-belief complex.” This just means that the individual’s extreme emotions produce two effects: first, they color her views of the future, leading her to adopt false beliefs about what her future life will most likely be like and second, while the emotion lasts, these beliefs are unchangeable, completely insensitive to counter evidence. Even assuming this is the right characterization, however, it is unlikely that the current framework would count her as lacking capacity because the primary effect of her grim view of the future is *a change in what she currently cares about and/or wants*. The current framework doesn’t allow us to question what someone wants or sees as important, but only allows us to identify formal flaws in the way such values or concerns feed into decision-making. Given Donna’s current preference for not having to face her grief, her choice is instrumentally rational.

Cases like these are genuinely possible and occur, though it is hard to say how common they are. Many patients with anorexia lack the insight of Terence and deny they are really ill or in danger (Tan, Stewart, Fitzpatrick & Hope 2003, 2006a). Likewise, many people emotionally overwhelmed by bad news lack insight into the fact that the way things look from their current standpoint is probably temporary. If lack of insight can be demonstrated, then such patients may count as incompetent in virtue of failing the appreciation requirement.⁶ However, there remain cases like Terence and Donna who have appreciation as generally defined, but who, intuitively, should not count as competent to decide.

At this point, it is important to highlight some of the background ethical constraints that have shaped bioethical thinking in this area. Though these constraints are not articulated as part of the four-abilities model itself, they are nonetheless often discussed along with it. Moreover, the fact that the model respects these constraints further explains its appeal.

⁶ For an example of the appeal to appreciation see the commentary on Tan et al. (2006a) by Grisso & Appelbaum (2006).

The first ethical constraint is commonly called “value neutrality.” A major ethical aim of modern medicine has been to ensure that competent patients are free to act on their own values even if those values differ from those of clinicians and even if they are highly unusual.⁷ To ensure that the freedom to act on one’s own values is real, many accept as a foundational principle of capacity assessment that capacity should never be determined *simply on the basis of what the patient chooses or wants* no matter how unusual this is. One way to think of it is this: A major lesson of the patient’s rights movement is that in some circumstances even death can be a rational choice. Once we grant this, we cannot say a person lacks capacity simply because she chooses death or a treatment path leading to death. Instead—the thinking goes—we must distinguish competent from incompetent choice by looking at the *process* that led up to the choice (Brock & Buchanan 1990, 50-51, 58).

A second important constraint is diagnostic neutrality. Just as capacity is not supposed to be determined simply by what the patient chooses, neither is capacity supposed to be determined simply on the basis of a diagnosis (Kim 2010, 11; Kim 2016, 189). This is particularly important for those with mental illness or cognitive deficits, since historically such individuals were generally viewed as incapable of making *any* decisions (Kim 2010, 11). The current framework allows that *some* individuals with mental illness may be globally incompetent, but insists that others may be globally *competent* and still others partially competent—able to make some choices but not others. If you believe that a patient with mental illness lacks capacity to make a particular decision, then (so the argument goes) what needs to be shown is not that they have such an illness, but that the illness is, in this very case, undermining the processes key to decision-making. On the standard model, this means showing that mental illness has undermined one or more of the four abilities.

The third constraint I refer to as “inclusivity.” It is widely accepted, though not often

⁷ The language of value neutrality is common. See e.g. Holroyd 2012, Kim 2016, as well as a full issue of the *International Journal of Law in Context*, devoted to value neutrality in competence assessment (Craigie 2013). The UK Mental Competence Act of 2005 states “a person is not to be treated as unable to make a decision merely because he makes an unwise decision.” (Section 1, Principle 4).

articulated. When building a model for assessing capacity, one must be careful not to build in too much, since then one might arrive at a model according to which too many people lack capacity. To avoid setting the bar too high, we must be careful not to rely on idealized notions of decision-making (Whiting 2015, 184-86). Ordinarily we assume that the majority of adult human beings are competent. And we need this to turn out to be true for ethical reasons. A major aim of the patient's rights movement of the 60s, 70s, and 80s was to ensure that most people be allowed to make their own health-related decisions. So, the goal is to minimize interference to the extent we can (Berg et. al. 1996, 377), which translates into ensuring that *most people are competent*. One result of this minimalist requirement, however, is that we must accept that even those who have capacity, can and sometimes do make mistakes, where this simply means they make choices regrettable from their own perspective. The freedom to make one's own choice is the freedom to make good as well as bad choices (Berg et. al 1996, 377).

The four abilities model has been so influential in part because it fits so easily within these constraints. It is focused on process as opposed to outcome. It makes no reference to either the patient's values or the patient's diagnosis. And its minimalist cognitive requirements ensure that most people turn out to be competent most of the time.

These constraints illuminate further why Terence and Donna are deemed to have capacity despite the fact that, intuitively, they lack it. In the second case, overwhelming emotions in response to major life-altering news seem to have made Donna unable to think about, or give weight to, her likely future. But that is presumably temporary. However, we cannot just say that she lacks capacity *because emotion is shaping her decision*. A broad appeal to emotion would rule too many people incompetent. Moreover, emotions are not always bad forces. They play a role in most of our decisions, including many of the best ones. Could we say instead that she is depressed, or traumatized? No, for while one or both claims *might* be true, using such facts to declare someone incompetent violates diagnostic neutrality.

Consider now Terence. In his case, the real stumbling block is his claim that he simply prefers to die rather than gain weight. He values thinness more than life itself. But as we saw above we cannot rule someone incompetent simply on the basis of unusual values or choices.⁸ That would violate value neutrality. Nor can we simply appeal to his anorexia, since this would violate the commitment to diagnostic neutrality. We do not want to say that all patients with anorexia are globally incompetent. Many who lack the capacity to make *some* decisions retain the capacity to make others.

§2.0 Diagnosing the Problem and Re-Conceptualizing Capacity

Given the way it developed, it is not really surprising that the current framework cannot capture all the cases it ought to, because it was intentionally set up to focus on *formal* features of decision-making. In their foundational work on capacity, Buchanan and Brock (1990, 18) note that to be competent is to be competent *to do* something, to perform an action with a certain level or degree of skill. In this case, we are concerned with the ability to make a decision of a certain quality, i.e. a good, or at least a “good enough” decision. However, Buchanan and Brock understand good decision-making in terms of formally good decision-making. On their account, to have capacity is to be able to make decisions that have enough of the marks of formally good decisions, i.e. they are coherent, instrumentally rational, etc.

However, it is important to remember why we value formally good decisions. When someone can’t put information together in a coherent way, or can’t follow very basic forms of instrumental reasoning, they are unlikely to make a decision that furthers the very end they intended to further with their decision. Even though formal properties matter, they matter *because* of the way they support a given end. But what is the end of decision-making in the kinds of cases

⁸ Tan et al. (2006a) suggest this might be permissible if the values are “pathological values,” i.e. values that derive from the mental disorder. However, this proposal faces various other problems that I cannot, for reasons of space, discuss here. For further discussion of this proposal see Vollman 2006; Tan et. al. 2006b; Whiting 2009; Tan et. al. 2009; Kim 2016; Hawkins & Charland 2020).

that concern us?

A natural answer emerges if we consider why we need to define decision-making capacity at all. It is often noted that the need to assess capacity arises from a tension between two values: well-being on the one hand, and free choice on the other (Buchanan and Brock 1990, 40-41). Any person could on occasion make a deeply foolish choice. Thus, the only way to completely protect well-being would be to severely limit free choice, which is unacceptable. On the flip side, complete embrace of free choice without any consideration of well-being would be disastrous for certain vulnerable, decisionally-impaired individuals, and so equally unacceptable. What we seek is the right balance. If we combine this thought with reflection on the third ethical constraint mentioned earlier, namely, that an acceptable account of capacity must entail that most people most of the time have capacity, the following natural answer emerges. The individuals we need to protect are those *who are less able to look after their own interests than most ordinary people are, individuals who are more likely than others to make serious prudential mistakes.*

This, in turn suggests that we frame the goal of medical decision-making in welfare terms. The general aim of medical-decision making is the patient's welfare. To have decision-making capacity is to be able to look after one's own interests *at least as well as most other people can.* And the goal of capacity assessment is to determine whether a given individual can, in fact, look out for her own interests at least this well.

However, many will object that welfare is not always the goal of patient medical decision-making. Sometimes patients choose less good care in order to save money or to ensure that their care is not too burdensome on loved ones. In a non-paternalistic world, competent adult patients are (and should be) free to choose in these ways if they wish. However, while this is true, it doesn't undermine my claim properly understood. Even though other considerations often come in, it is common to think of medical care as focused on the patient's good. It seems fair to say that the patient's good is the *default* goal of such decisions, the goal we assume unless made aware that

other concerns are at stake. And because of this, I think it is also fair, when we come to think about capacity assessment, to insist that individuals *be able* to look after their own welfare at least as well as most other people can. If you are able to look out for your interests to this degree, then you should be free to decide things for yourself, including deciding against your own interests if that's what you want. But if you are not even able to look out for yourself to this degree, then you should not be given the freedom to cast your own welfare aside, since that kind of freedom is not really meaningful.

This way of understanding capacity assessment also suggests a diagnosis of the problem cases. What bothers us in Terence's case and in Donna's is our sense that these individuals are making decisions that run counter to their own interests, subjectively construed (more on what that means below). Their choices are not bad in a formal sense, but bad relative to the presumptive aim of decision-making. Of course, merely making a poor choice is not by itself (and never should be) sufficient to establish that someone lacks decision-making capacity. But here is where the reconceptualization can helpfully diagnose our sense that both lack decision-making capacity. Not only do we feel certain that their decisions go against their interests, we also feel certain that these particular choices are *no mere accident*. Making a poor choice when one is in shock is both predictable and seems to come about because of the way situations like that affect the mind, namely, they make a person incapable of thinking about and/or giving weight to their likely future. Likewise, in the context of anorexia nervosa refusing life-sustaining treatment that involves weight gain is both predictable and the result of the rigid evaluative outlook characteristic of the disorder. Both patients therefore strike us not only as people who choose poorly, but as people who we suspect from the outset are more likely to choose poorly than ordinary people.

Of course, many will have the reaction that, even if I'm correct to this point, there is no ethically sound way to alter our practices to bring these thoughts to bear on real life cases. But as I now hope to illustrate, there might be.

§3.0 A New Proposal for Difficult Cases

I propose that we continue to consider the traditional four abilities as *necessary* for capacity. The suggestion that follows is not a replacement, but an add-on. With that in mind, I propose that even in cases where individuals have all four abilities, if and only if *both* of two further requirements are satisfied, the patient should be deemed incompetent to make the specific decision at hand. The two requirements are:

(1) There must be good evidence for thinking that the patient is making a serious, prudential mistake here and now, and

(2) The patient must be known to have a condition that, in turn, is known to make those who have it more likely than ordinary to make prudential mistakes.

Each requirement, were it employed by itself, would be ethically problematic. However, when both are satisfied there is, I submit, nothing ethically problematic at all.

§3.1 The First Requirement

Requirement (1) says we must have good evidence of current prudential mistake. Thus, to determine whether it is satisfied we need both an account of welfare and an account of serious prudential mistake. Objectors may press that we don't currently have a theory of welfare—at least, not one widely accepted. Moreover, if we must wait for philosophers to agree on the true theory, we may have to wait for eternity! If, on the other hand, we simply rely on ordinary ideas about welfare, such as the idea that it is generally better to preserve life, or generally better to avoid disability, then we risk imposing significant values on individuals for whom those values are inappropriate.

Despite these legitimate concerns, the proposal can be made to work without settling on anything as controversial as a full theory of welfare. Instead, I propose to appeal to three broad *components* of welfare that almost *any* theory will recognize as having weight. We should consider on the positive side, (1) happiness, understood *not* as pleasure or joy or any such fleeting experience, but as a generally positive, emotionally-grounded, outlook on life. We should also consider (2) evaluative engagement, by which I mean a person's direct engagement with people and projects that matter deeply to her. Evaluative engagement might mean participating in relationships or engaging with projects one cares about or working towards valued goals. Most theorists and lay people agree that, other things being equal, people are better off when they are happy *and* better off when they are able to engage in the right ways with the things (and people) that matter most to them.

On the negative side we should consider (3) all forms of suffering, which includes extreme physical pain as well as all forms of emotional or psychological suffering. Almost everyone can agree that suffering is both intrinsically bad and instrumentally bad (it undermines a person's ability to engage with projects and people that matter to her).

In addition, these elements must be interpreted *subjectively*. In other words, we want to ask whether, depending on the option chosen, this individual is likely to suffer or be happy or be able to pursue what matters to her in a way that she sees as significant.

Finally, and importantly, the question being asked is whether an individual seems to be making a *serious prudential mistake*. The qualifier, "serious," matters, since it would require a much more fine-grained (and more controversial) theory of welfare to detect small prudential mistakes—for example, choices that are subjectively bad for a person, but only a little bit more so than some other choice on offer. All we are concerned with are *serious prudential mistakes*—cases in which a person is about to irrevocably choose something much, much worse for her than something else easily available.

Here are two examples to illustrate: (1) It would be a serious prudential mistake to irrevocably choose something that leads to overall suffering (a future with significantly more negatives than positives) when that could have been easily avoided; and (2), it is a serious prudential mistake for a subject to choose death in cases where, if the subject lived, her life would contain significantly more positives than negatives.

To decide whether or not a patient is making a serious prudential mistake one must try to consider what the most likely outcomes of each potential choice would be, in rough terms. Would the life be dominated by suffering? Would it have as many opportunities for pursuit of her values? Is she likely to be able to find happiness in that life? Good answers will require knowledge of the particular individual, her values, her likes and dislikes, as well as her psychological dispositions. In cases involving mental illness it is relevant how likely the subject is to improve or recover and what her quality of life would be like if she did versus if she did not.

Importantly, I do not imagine asking untrained individuals or family members to make such assessments. Rather, I imagine that *if* this idea were accepted, it would be necessary to train certain healthcare professionals to think in terms of the above welfare elements, giving them much more concrete articulations of each. No doubt many people would be involved in *developing* such training materials. And it would also probably be advisable to develop instruments to guide conversation with patients about their values and what makes them happy and so on. I do not imagine that this is the kind of assessment that would be left to general clinicians, since it is simply not likely that all clinicians could receive the proper training. But since the proposal only applies in special cases, I imagine that competence would be determined by a psychiatric consultant familiar with the patient group in question, someone who has had the relevant training, knows how to apply (1) and (2) together, and who understands that the ethicality of the policy *depends* on the subjective construal of welfare. Finally, for various reasons it seems safest to require that several people be involved in any particular assessment decision that makes use of these requirements and that these people be

required to justify their assessment to one another in the terms of the model. If no agreement is possible about whether or not a serious prudential mistake is in the making, the first requirement is not satisfied and so the patient should, assuming she has the initial four abilities, count as competent to make her own choice.

These practical requirements (which could be refined) serve the purpose of ensuring that not just any ideas about patient welfare are allowed to influence determinations of competence. In particular, it is to ensure that clinicians do not assume that decisions that fail to promote certain medical values are thereby contrary to an individual's welfare, and equally to ensure that family members and friends do not assume that what would be good for them in circumstances like the patient's is therefore good for the patient. I am well aware of the disturbing history of making decisions for others that do not actually promote their welfare, but instead promote someone else's vision of welfare. I would not endorse any use of these two additional requirements if I did not believe adequate safeguards could be put in place to keep that from happening.

As an example, consider once again anorexia nervosa. I do not suppose that all patients with anorexia who refuse life-sustaining treatment involving weight gain are making a serious prudential mistake. In many cases, it is not at all clear that they are. For many patients with severe and enduring anorexia nervosa, the chances of full recovery are very low, and a significant number of these patients report low quality of life (Treasure et. al. 2015). To be clear, I am not saying such a choice (to refuse treatment) would be good, but only that it is not crystal clear that it would be bad. And if it is not clear that the individual is making a seriously poor choice subjectively construed, then requirement (1) is not met and the patient counts, on my view, as competent.

The difference in the case of Terence is that he does not (yet) have severe and enduring anorexia nervosa. His chances of full recovery with treatment are still high (Treasure, et. al. 2015).⁹

⁹ Any claims made in this article about the verdicts my model would give in particular cases are simply guesses shared for the purpose of illustrating how the model is intended to work. Real verdicts would, of

It is against that background that his choice looks like a serious prudential mistake because it seems that he is throwing away many years of life that would be quite good for him *from his perspective if he lived them*.

Similarly it seems plausible to suppose that Donna is making a serious prudential mistake because she has years of life ahead, many good friends, and an interesting career, all of which suggest that she herself would find meaning and happiness again, once her shock and grief subside. She is currently, understandably distraught, but it seems relevant that she has recovered from depression before. Perhaps our judgment would be different if she had a life-long history of struggling with treatment-resistant depression, for then it would be *less* clear whether she would fully recover. However, given that most people eventually recover from the breakup of a marriage, and given that she has much to live for (her art, her friends), it is plausible to suppose she is making a serious prudential mistake.

Because the proposal is so at odds with the way many have been trained to think about capacity assessment, it can seem as if it *must* be unethical. I fully grant that if (1) were used by itself that would be ethically problematic. But the proposal requires (1) and (2) together. Still, some may feel that, even taken together, my two requirements are problematically paternalistic and/or in deep tension with the requirement of value neutrality. Let us consider each in turn.

First, it is important to note that “paternalism” is not the right label for the concern many feel. Recall that, as most theorists use the term, not all paternalistic actions—i.e. actions taken against a person’s will for that person’s benefit—are morally problematic. For example, it is not morally objectionable to treat a small child paternalistically. What is objectionable is treating a *competent adult* paternalistically (Feinberg 1986, 3-8). In other words, we use competence as the dividing line such that we only know we have a case of morally objectionable paternalism *if* it is

course, depend on careful specification of the central concepts sketched here as well as sufficient amounts of high quality evidence to support relevant claims.

directed at a competent adult. Given that we are currently trying to decide who counts as competent, labeling the proposal “paternalistic” begs important questions.

However, in those cases where paternalism is morally objectionable this is because it infringes on an important kind of freedom, namely, freedom of choice in the self-regarding sphere. Although one might worry that my proposal unduly limits this kind of freedom, closer examination reveals that it doesn't. Recall that the freedom to make choices about your own life is a freedom that is not supposed to be constrained by whether you make good choices. It includes the freedom to make prudential mistakes, including disastrous ones. However, as noted earlier, it is also generally agreed that an acceptable theory of capacity must have the result that most people most of the time are competent. It is as if, as a society, we have agreed that ordinary adults have a degree of decision-making capacity that, while not perfect, is *good enough*. But then, given that my proposal uses this as a significant marker, it will not unduly infringe on personal freedom of choice. The need to restrict the scope of requirement (1) in precisely this way is what requirement (2) is for.

On the other hand, this proposal violates the requirement of value neutrality *as this is usually understood* because it allows us to occasionally consider the welfare impact of a choice made in light of a person's current values and preferences.¹⁰ But although the proposal conflicts with the usual formulation of this rule, it does not, I submit, conflict with the ethical concerns that motivated adoption of the rule.

As I understand it, value neutrality serves two ethical purposes in the current framework. Its first important function is to block the imposition on patients of values that do not promote their welfare *subjectively construed*. So, for example, it blocks clinicians from judging that a patient is making a poor choice because she chooses contrary to favored medical values. It likewise blocks

¹⁰ Pickering et al. (2022) argue that “in some cases [a person should] be judged incapable of making [a] decision *because of the harmfulness of the decision*.” This proposal is problematic for the same reasons my requirement (1) *would be problematic on its own*. We need to limit appeals to the welfare outcome of a choice and we need a justification for why the limit is what it is. Otherwise we risk impinging on personal freedom too much.

other kinds of judgments to the effect that a person is making a poor choice, simply because she chooses something that others see as odd or would not choose for themselves. Third parties are often tempted to judge choices in these ways, but this generally leads to decisions that are not good for the individual subjectively construed. Thus, the first function of value neutrality is to ensure that values that conflict with patient welfare subjectively construed are not allowed to drive determinations of competence.

What confuses many people is that I do not simply equate a person's welfare subjectively construed with the realization of an individual's current values and preferences. However, when we think about it, this is not really so odd. Values and preferences at a given time are usually *relevant* to welfare, but are never guaranteed to be a foolproof guide to it. After all, it is a familiar fact that people often adopt values at one time only to find later that the pursuit of those same values undermines either their happiness or their pursuit of other values that are even more important to them. In such cases, individuals typically *view themselves as having made a poor choice*. Thus, what I am here suggesting is that the proper way to assess the subjective welfare impact of a choice is to consider how the decision and its consequences will most likely be viewed by the individual over time as she lives with those consequences. The relevant questions to ask are questions about how much *she* will enjoy a given future, how much meaning *she* will likely find in a given life and so on. In such cases, what we are doing is trying to make limited judgments (insofar as these are possible) about the degree to which a person's *current desires or values fit with her own welfare as she will herself ultimately come to view it*.

The second important moral function of value neutrality is linked to the idea that people should be free to make mistakes. I have already explained why my proposal would not unduly limit such freedom. But it is useful to see how the commitment to value neutrality is linked to concern for freedom. Given that it is not the role of capacity assessors to ensure that individuals always pursue what is best for them, one might then suppose that there could *never* be a good or legitimate reason

for considering the welfare impact of a patient's choice in the process of assessing capacity. And if there is no such need, then (so the thought goes) one might as well insist that no consideration of the welfare impact of choice occur. But what I have been trying to argue here is that while it is true that the welfare impact of a choice should never by itself justify a finding of incompetence, it is not true that there could never be a legitimate reason for considering the welfare impact of a choice when assessing capacity. While it is usually not relevant, it may in a small number of cases be relevant.

The requirement of value neutrality, I submit, is based partly in a deep commitment to personal freedom, and partly in a deep commitment to the idea that welfare is individually relative and subjective. I share these commitments, and my proposal honors them.

§3.2 The Second Requirement

Now consider requirement (2), which is also necessary. Requirement (2) appeals to the idea of a "condition" known to increase the likelihood of prudential mistakes. To understand my use of "condition," it is necessary to understand the work the requirement is meant to do. We need evidence not only that a given person is choosing badly, *but that her decision is different from ordinary sorts of decisions that just happen to be unwise*. My idea is that we can obtain support for such a claim—the claim that an individual's decision reflects an inability to track her own interests at least as well as most people can—by appealing to what is known about the decision-making patterns of a particular, clinically defined group to which she belongs. In most cases, the group would likely be defined by a common diagnosis such as e.g. "anorexia nervosa." But I use the word "condition" rather than "mental disorder, or even just "diagnosis" because I want to leave room for the idea that there might be some clinically defined, relevant groups with no agreed upon diagnostic label.

For example, Donna fits the layperson's notion of "being in shock." However, in discussions with different psychiatrists, I have received different answers about whether Donna fits any specific diagnosis. For example, although she currently sees everything in negative terms, she does not count as clinically depressed since that diagnosis partly requires that symptoms persist for a while. However, even though no diagnosis can be agreed upon, there is a distinctive feature of the cases I have in mind. Namely, a certain sub-set of people, in the immediate aftermath of life-changing news or events, are unable to think or care about their most likely future. I am not the first to suggest that this inability may be relevant to capacity assessment.¹¹ However, this inability is not a formal flaw in decision-making, but something that matters *because it predictably leads a person to make poor welfare choices*. So it is only once we accept that the aim of assessing decision-making capacity is to determine whether individuals are as able as most people to look out for their own interests that we can explain why this is a clinically relevant group.

In other cases, it seems to me that the relevant group might turn out to be a specific sub-set of those with a particular diagnosis, e.g. those who have had an anorexia diagnosis for less than x number of years. However, and this is the most important point, it is not my intention that "condition" be left vague. For one thing, members of a group who share a "condition" would need to have more in common than just a shared tendency to make certain kinds of poor decisions. There would need to be other situational and clinical factors uniting them. And ultimately, if including groups without a clear diagnosis creates problems, I would simply recommend focusing more narrowly on clear diagnostic groups.

Of course, there is more to understanding (2) than just understanding what a condition is. In order to make use of this approach ethically we would need to acquire relevant evidence. We would need to establish that members of the relevant group do in fact make serious prudential

¹¹ I have already mentioned Halpern (2001; 2011, 2012). See also Craigie (2013b) who explores the relevance of the ability to accurately imagine different possible futures.

mistakes more often than normal. Nor would this always be straightforward. For example, even though it is true that patients with anorexia nervosa have a clear pattern of making choices that endanger their lives and which often lead to their deaths, this is not enough. For we can't simply assume that a choice that predictably leads to death is a serious prudential mistake. However, we could study treatment outcomes for patients with anorexia nervosa, and collect data over time on the quality of life of patients who fully recover and patients who do not. Assuming we found that the quality of life of patients who fully recover is generally good, this would support the idea that choosing death when one's chances of full recovery are still high counts as a serious prudential mistake.

Finally, it is important to note that even when the condition appealed to is a mental disorder, the particular way (2) functions ensures that we are not in conflict with the ethical concerns that support diagnostic neutrality. Importantly, on this proposal, it is never the case that a diagnosis alone is used to deem someone incompetent. Moreover, the proposal ensures that the individual is left free to make other decisions, and that people with the diagnosis in question are not viewed as generally lacking in capacity.

§4.0 Conclusion

The current framework for assessing decision-making capacity—the four abilities model—cannot account for various cases in which people make self-destructive choices that are not simply random bad luck. In response, I have argued that we should understand decision-making capacity as *the ability to track one's own personal interests at least as well as most people can*. I have then suggested that we should treat the familiar four abilities as necessary but not always sufficient for decision-making capacity. I have proposed two new, additional requirements that would only come into play in a handful of cases. In those cases, however, when *both* requirements are satisfied, we can justify the claim that the individual choosing currently lacks the capacity to make the particular decision at

hand. Although each of the two requirements would be objectionable on its own, together they offer a perfectly ethical addition to the four abilities framework—one that could greatly improve capacity assessment in certain kinds of difficult cases.

To implement my proposal would require the development of clear guidelines about how welfare interests are to be understood and applied. The proposal would need to be backed by research about the decision-making patterns of particular clinical groups to whom the proposal might apply. And finally, there would need to be various safeguards in place to ensure the requirements are used as intended. If this were all done, I believe use of this proposal would be ethical. Of course, one might continue to have reservations about a policy that, while ethical in principal, might not be used ethically in practice, especially if one fears that adequate safeguards could not or would not be developed. Trying to decide how well-based such worries are would be its own project. At this point, I simply hope enough readers see the promise in the approach to think it worth exploring further.¹²

References:

- Appelbaum, P. S., T. Grisso. 1988. Assessing patients' capacities to consent to treatment. *New England Journal of Medicine* 319 (25): 1635-38.
- Appelbaum, P. S., T. Grisso. 1995. The MacArthur Treatment Competence Study I: Mental illness and competence to consent to treatment. *Law and Human Behavior* 19 (2):105-26.
- Appelbaum, P. S., T. Grisso. 2001. *MacArthur competence assessment tool for clinical research (MacCAT-CR)*. Sarasota, FL: Professional Resource Press.

¹² Thanks to audiences that have given helpful feedback, including attendees of Psychiatry Grand Rounds at Duke Medical School, and attendees of talks given at/for The University of Cape Town Medical School, The UNC Philosophy and Psychiatry Research Group, The Oxford-Uehiro Centre for Practical Ethics, The Center for Bioethics and Medical Humanities of The Medical College of Wisconsin, the Kansas Philosophy of Well-Being Workshop, and the Rocky Mountain Ethics Congress. Special thanks to Laura Weisberg, Kathryn Petrozzo, Laura Specker Sullivan, Roger Crisp, and Walter Sinnott-Armstrong. I dedicate this paper to my friend and co-author Louis Charland, who, sadly, passed away unexpectedly in 2021. He would be happy to see this in print. Finally, huge thanks to my friend Jodi Halpern, who first sparked my interest in capacity, and with whom, over the years, I have had many fruitful discussions of these matters.

- Banner, N. F. 2013. Can procedural and substantive elements of decision-making be reconciled in assessments of mental capacity? *International Journal of Law in Context* 9 (1): 71-86. [Special Issue on Value Neutrality]
- Berg, J. W., P. S. Appelbaum, T. Grisso. 1996. Constructing competence: Formulating standards of legal competence to make medical decisions. *Rutgers Law Review* 48 (2): 345-396.
- Breden, T. M., J. Vollmann. 2004. The cognitive based approach of capacity assessment in psychiatry: a philosophical critique of the MacCAT-T. *Health Care Analysis* 12 (4): 273-83.
- Buchanan, A. E., D. W. Brock. 1990. *Deciding for others: The ethics of surrogate decision-making*. Cambridge: Cambridge University Press.
- Bursztjn, H. J., H. P. Harding, T. G. Gutheil, A. Brodsky. 1991. Beyond cognition: The role of disordered affective states in impairing competence to consent to treatment. *Journal of the American Academy of Psychiatry and Law* 19 (4): 367-381.
- Charland, L. C. 1998a. Appreciation and emotion: Theoretical reflections on the MacArthur Treatment Competency Study. *Kennedy Institute of Ethics Journal* 8 (4): 359-77.
- Charland, L. C. 1998b. Is Mr. Spock mentally competent? Competence to consent and emotion. *Philosophy, Psychiatry, Psychology* 5 (1): 67-95.
- Charland, L. C. 2006. Anorexia and the MacCAT-T Test for mental competence: validity, value, emotion. *Philosophy, Psychiatry, Psychology* 13 (4): 283-7. [Commentary on Tan et al. 2006a.]
- Craigie, J. 2013a. Introduction: Mental capacity and value neutrality. *International Journal of Law in Context* 9 (1): 1-3. [Special Issue on Value Neutrality]
- Craigie, J. 2013b. Capacity, value neutrality and the ability to consider the future. *International Journal of Law in Context* 9 (1): 4-19. [Special Issue on Value Neutrality]
- Elliot, C. 1997. Caring about risks: Are severely depressed patients competent to consent to research? *Archives of General Psychiatry* 54 (22): 113-16.
- Faden, R., T. Beauchamp. 1986. *A history and theory of informed consent*. New York: Oxford University Press.
- Feinberg, J. 1986. *Harm to self: Moral limits of the criminal law, volume 3*. New York: Oxford University Press.

- Freyenhagen, F., T. O'Shea. 2013. Hidden substance: Mental disorder as a challenge to normatively neutral accounts of autonomy. *International Journal of Law in Context* 9 (1): 53-70. [Special Issue on Value Neutrality]
- Grisso, T., P. S. Appelbaum. 1995. The MacArthur Treatment Competence Study. III: Abilities of patients to consent to psychiatric and medical treatments. *Law and Human Behavior* 19 (2): 149-174.
- Grisso, T., P. S. Appelbaum. 1998a. *The assessment of decision-making capacity: A guide for physicians and other health professionals*. New York: Oxford University Press.
- Grisso, T., P. S. Appelbaum. 1998b. *MacArthur competence assessment tool for treatment (MacCAT-T)*. Sarasota, FL: Professional Resource Press.
- Grisso, T., P. S. Appelbaum. 2006. Appreciating anorexia: Decisional capacity and the role of values. *Philosophy, Psychiatry, & Psychology* 13 (4): 293-297.
- Grisso, T., P. S. Appelbaum, C. Hill-Fotouhi. 1997. The MacCAT-T: A clinical tool to assess patients' capacities to make treatment decisions. *Psychiatric Services* 48 (11): 1415-19.
- Grisso, T., P. S. Appelbaum, E. P. Mulvey, K. Fletcher. 1995. The MacArthur Treatment Competence Study. II: Measures of abilities related to competence to consent to treatment. *Law and Human Behavior* 19 (2): 127-148.
- Halpern, J. 2001. *From detached concern to empathy*. New York: Oxford University Press.
- Halpern, J. 2011. Emotions, autonomy, and decision-making capacity. *AJOB Neuroscience* 2 (3): 62-3.
- Halpern, J. 2012. When concretized emotion-belief complexes derail decision-making capacity. *Bioethics* 26 (2): 108-16.
- Hawkins, J. Considering the welfare impact of a choice when assessing capacity: Is it always wrong?" In: C. Carrozzo and E. Ritchie (eds.) *Mental capacity: Psychiatric, psychological, and philosophical perspectives* (Oxford: Oxford University Press), forthcoming.
- Hawkins, J., L. C. Charland. 2020. Decision-making capacity. *The Stanford Encyclopedia of Philosophy* (Fall 2020 Edition), E. N. Zalta (ed.), URL = <<https://plato.stanford.edu/archives/fall2020/entries/decision-capacity/>>.
- Hermann, H., M. Trachsel, B. S. Elger, N. Biller-Andorno. 2016. Emotion and value in the evaluation of medical decision-making capacity: A narrative review of arguments." *Frontiers in Psychology* 7: 765.

Holroyd, J. 2012. Clarifying capacity: Value and reasons. Chap 7 in *Autonomy and mental disorder* (ed.) L. Radoilska. New York: Oxford University Press.

Hope, T., J. Tan, A. Stewart, J. McMillan. 2013. Agency, ambivalence and authenticity: The many ways in which anorexia nervosa can affect autonomy. *International Journal of Law in Context* 9 (1): 20-36. [Special Issue on Value Neutrality]

Kim, S. Y. H. 2010. *Evaluation of capacity to consent to treatment and research*. New York: Oxford University Press.

Kim, S. Y. H. 2016. The place of ability to value in the evaluation of decision-making capacity. Chap 7 in *Philosophy and Psychiatry: Problems, Intersections and New Perspectives* (eds.) D. Moseley and G. Gala. New York: Routledge.

Mackenzie, C., W. Rogers. 2013. Autonomy, vulnerability, and capacity: A philosophical appraisal of the Mental Capacity Act." *International Journal of Law in Context* 9 (1): 37-52. [Special Issue on Value Neutrality]

Mental Capacity Act 2005: <https://www.legislation.gov.uk/ukpga/2005/9/contents>.

Pickering, N., G. Newton-Howes, G. Young. 2022. Harmful choices, the case of C, and decision-making competence. *The American Journal of Bioethics* 22 (10): 38-50.

Richardson, G. 2013. Mental capacity in the shadow of suicide: What can the law do? *International Journal of Law in Context* 9 (1): 87-105. [Special Issue on Value Neutrality]

Tan, J. O. A., T. Hope, A. Stewart. 2003. Competence to refuse treatment in anorexia nervosa. *International Journal of Law and Psychiatry* 26: 697-707.

Tan, J. O. A., A. Stewart, R. Fitzpatrick, R. A. Hope. 2006a. Competence to make treatment decisions in anorexia nervosa: Thinking processes and values. *Philosophy, Psychiatry, and Psychology* 13 (4): 267-82.

Tan, J. O. A., A. Stewart, R. Fitzpatrick, R. A. Hope. 2006b. Studying penguins to understand birds. *Philosophy, Psychiatry, and Psychology* 13 (4): 299-301.

Tan, J. O. A., A. Stewart, T. Hope. 2010. Decision-making as a broader concept. *Philosophy, Psychiatry, and Psychology* 16 (4): 345-9.

Treasure, J., D. Stein, S. Maguire. 2015. Has the time come for a staging model to map the course of eating disorders from high risk to severe enduring illness? An examination of the evidence. *Early Intervention in Psychiatry* 9 (3):173-184.

Jennifer Hawkins

Pre-print of Target Article American Journal of Bioethics, Published Online June 26, 2023

DOI: 10.1080/15265161.2023.2224273

- United Nations General Assembly. 2006. Convention on the Rights of Persons with Disabilities: <https://www.un.org/development/desa/disabilities/convention-on-the-rights-of-persons-with-disabilities/article-12-equal-recognition-before-the-law.html>.
- Vollmann, J. 2006. 'But I don't feel it': Values and emotions in the assessment of competence in patients with anorexia nervosa. *Philosophy, Psychiatry, and Psychology* 13 (4): 289-91. [Commentary on Tan et. al. 2006a].
- Weltzen, T. E. 2016. Treating males with severe and enduring anorexia nervosa: Different to females? Chap 15 in *Managing Severe and Enduring Anorexia Nervosa: A Clinicians Guide* (eds.) S. Touyz, D. Le Grange, J. H. Lacey, P. Hay. New York: Routledge.
- White, B. C. 1994. *Competence to consent*. Washington, D.C.: Georgetown University Press.
- Whiting, D. 2009. Does decision-making capacity require the absence of pathological values? *Philosophy, Psychiatry, Psychology* 16 (4): 342-4.
- Whiting, D. 2015. Evaluating medico-legal decisional competency criteria. *Health Care Analysis* 23 (2): 181-96.