

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 widely accepted ethical constraints on decision-making assessment.

Introduction

In our society, people value control over their lives. They value making their own medical decisions. Because of this we have reached a point where patient authority over treatment decisions is a default assumption. Informed consent—an important part of what helped to make patient authority real—has now 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.¹ 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.

In biomedicine 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

¹ In a recent article also discussing capacity assessment, Pickering et al. (2022) note that two key assumptions of their project—(1) that some people lack the capacity to make certain decisions and (2) that in such cases substitute decision-making of some form is ethically justified—are rejected by the United Nations Convention on the Rights of Persons with Disabilities (Article 12, 2006). I also accept (1) and (2), disagreeing with the UNCRPD on these two points. However, once the meaning of various terms and the actual nature of certain practices is clarified, I believe the degree of disagreement between myself and the UNCRDP is much smaller than it initially appears. Ultimately, I suspect the degree of disagreement is quite small. Unfortunately, however, defending this is beyond the scope of the present article and must await another occasion.

“capacity”) or whether someone is mentally competent (often just shortened to “competent”).² 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 somewhat surprising to discover 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 great 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., the U.K., 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. I argue, however, that this is not the case.

There are certain kinds of cases (described below) that the framework seems unable to handle. However, to date no promising proposals have been made about how to change the

² 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).

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 are when it comes to assessing it. 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 need to 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 requirements. Without careful development of the ideas and collective agreement on interpretation of key elements, that would be deeply problematic. My goal is the more modest one of demonstrating the promise of this approach in the hope that others will agree with me 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.³ They have not only developed a framework listing four key abilities, but have also developed guidance for clinicians and researchers on how to interpret and assess these abilities

³ 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 two are close, and that both emphasize cognition in a similar way. Thus the concerns raised here plausible apply to it as well.

(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; Breden & Vollmann 2004; Tan, Stewart, Fitzpatrick & Hope 2006; Vollman 2006; Halpern 2011, 2012; Hermann 2016). In general, the critics have not doubted that the four abilities identified by the model 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 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 most 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. Terence has anorexia nervosa and it has been less than two years since he was diagnosed.⁴ Terence 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 go into hospital and allow himself to be fed to gain weight. But he refuses, saying that he knows

⁴ Although anorexia nervosa is more common among females, males make up approximately 10% of patients (Weltzen 2016).

he has an illness, he knows he is incredibly thin, and he knows that he may die. But 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.⁵

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 overall values he currently has and the priority he gives them.

Now consider a woman in her mid-50s I shall call Donna. 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 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

⁵ 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).

her change in attitude and the forcefulness of her demands her doctor calls for a psychiatric consult. The psych resident who comes to interview Donna discovers, after some probing, that Donna's husband visited her in the 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. 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.⁶

Like Terrence, Donna is deemed to have capacity because she understands her 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. Indeed, given her preference for not having to re-build her life without a spouse, 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 in shock or who are overwhelmed by grief 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.⁷ 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

⁶ The case of Donna is an altered version of the case "Ms. G" presented in Halpern (2001). I deliberately changed the original to remove elements that, in my experience, distract from the issue of decision-making capacity.

⁷ For an example of the appeal to appreciation see the commentary on Tan et al. (2006a) by Grisso & Appelbaum (2006).

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.

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 the values of clinicians and even if those values are highly unusual.⁸ To ensure that the freedom to act on one’s own values is real, many people 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. But 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

⁸ 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).

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 articulated as frequently as the first two. 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 first case, overwhelming emotions arising from shock or grief seem to have completely altered Donna’s outlook on life. 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.⁹ 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. As already noted, it was intentionally set up to focus on *process* without any reference to *outcome*. Moreover, process was to be assessed relative to whatever aim or desire the patient currently has. But this overlooks the complicated relationships people stand in to their desires and values.

It matters not only whether a person's choice makes sense in light of some current value or desire, but also whether the choice makes sense *given the goal or end that decision-making in this context should have*. Call this the "normative aim" for a type of decision-making. In cases where a type of decision-making is recognized as having a normative aim, we judge the goodness (or badness) of a decision by considering, in part, how it relates to the normative aim.

An example can help clarify. Jill is a manager trying to decide how best to handle a conflict that has arisen among her employees. As a *manager* it is her job to try to restore good will among her employees and thereby restore productivity. Of course, it is not always possible to please everyone. Success can be difficult, depending on the personalities involved. Still, *qua* manager, she

⁹ 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 too is problematic, since anorexia nervosa is partially defined by the tendency to embrace these values and/or the desire for thinness. For further discussion of this proposal see Vollman 2006; Tan et. al. 2006b; Whiting 2009; Tan et. al. 2009; Kim 2016; Hawkins & Charland 2020).

should look for a solution that is fair and which will satisfy most people. For the purpose of argument, let's grant this is a reasonable description of the decision-making goal a manager in this situation would and should have.

But suppose Jill is not a good manager. Maybe her own personal aims weigh more with her than being a good manager, and so she focuses on pleasing one person in particular (a friend who can do her a favor later on). My point is just this: Jill's decision might make sense given her actual goal at the moment (it might be instrumentally rational given her aim of appeasing the friend), but still not make sense given the normative aim of managing. Ordinarily, we would assess her decision as good or bad in terms of two features together: i.e. whether it aims at the appropriate goal, and whether it is instrumentally rational relative to the appropriate goal.

How then might this relate to medical decision-making? First, as illustrated above, in order to really assess whether someone is able to make a particular decision—at least in certain contexts, of which medicine is one—you have to have a prior sense of what the goal of that type of decision-making is.

In the particular case at hand, we are interested in understanding whether a person has the abilities she needs in order to do *well enough* at her task. As noted earlier, there is no requirement in medical contexts that people be *excellent* decision-makers. That would violate “inclusivity.” But even in order to say what “*well enough*” is and identify the necessary abilities we need an account of the relevant normative aim—an account of the normative aim of medical decision-making. I suggest we should think of the goal of medical decision-making *in welfare terms*. In other words, the goal of medical decision-making is that of identifying which medical option would best promote the patient's welfare. To count as having decision-making capacity, one must therefore be *good enough* at making decisions that track one's own welfare or interests.

No doubt some people will object that this is not always the goal of 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 free to choose in these ways if they wish.

However, while this is true, I do not think it undermines my claim. Even though other considerations often come in, it is common to think of medical care as naturally focused on the patient's good. Other concerns are seen as precisely that—*other* concerns, or side concerns. Thus, I think it is fair to say that patient welfare is normative for medical decision-making in the sense 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 interests 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. That kind of freedom is not really meaningful.

This, I suggest, is a much better way to conceptualize what we are trying to determine when we assess capacity. However, various people may assume that even if that's correct, 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 is.

§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 intended as a replacement, but as an add-on. With that in mind, I propose that even in cases where individuals have all four abilities, if *both* of two further requirements are satisfied, then 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 to make prudential mistakes than ordinary people.

Each requirement, were it employed by itself, would be ethically problematic for reasons I explain below. My claim, however, is that when both are satisfied they need not be ethically problematic at all. I will begin by explaining requirement (1) on its own—how it should be understood and what would need to be the case for it to be satisfied. Then I will turn to consider the ethicality of requirement (1) by itself, and how it fares in conjunction with requirement (2).

Requirement (1) requires that we 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. However, an objector may press, 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, for example, the idea that it is generally better to preserve life, or generally better to avoid disability, then we risk imposing significant values on individuals who do not share them.

Despite these legitimate concerns, this proposal can be made to work without settling on anything as complicated or 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. I suggest that we consider on the positive side, (1) happiness, understood not simply as pleasure or joy or any such fleeting experience, but as a generally positive, affectively-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 doing the

things one does to nurture and sustain a relationship: spending time with loved ones, helping them in various ways etc. It can also mean 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 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. Theorists of diverse types all 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 probably more controversial) theory of welfare to be able 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 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 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 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. Finally, in particular cases, it seems safest to require that several people be involved in any particular assessment decision that makes use of my requirements. 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.

Consider the fact that the chances for full recovery from anorexia nervosa are quite high in the early stages of the disorder but decline significantly over time. This means that in a case like Terence, whose anorexia nervosa is still in the early stages and for whom full recovery with treatment is both possible and likely¹⁰, starving himself to death now would be throwing away many years of life that could 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, despite her disability, she has years of life ahead, many good friends, and an interesting career suggesting that she herself would find meaning and happiness again, once her grief subsides. She is currently, understandably distraught, but she has recovered from depression before. Given that most people

¹⁰ 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 course, depend on careful specification of the central concepts sketched here as well as sufficient amounts of high quality evidence to support relevant claims.

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.

§4.0 But Is This Ethical?

One might suspect that the use of these requirements would be unethical. Admittedly, determining that someone who has the four abilities is incompetent simply on the basis of requirement (1) alone *would be* unethical. Likewise, it would be unethical to appeal to requirement (2) without (1). Each of the requirements is problematic if used *by itself*, but not when they are used together.

Requirement (1) might seem to judge the subject's values in an objectionable way, thereby violating value neutrality. However, that is not really correct. Recall that value neutrality requires that we never declare incompetence simply on the basis of what someone wants or desires or values, even if we think what they want is bad. It aims to rule out cases in which we impose our own sense of what is important on someone else, cases in which we fail to consider welfare subjectively—in terms of what matters to this person and what makes this person happy. The way I have described requirement (1) should make it clear that when we assess whether a person is making a serious prudential mistake we are not (at least not if we do it subjectively as we are supposed to) passing judgment on particular values. Rather we are making limited judgments 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*. It is a familiar fact that people sometimes adopt values at one time only to find later that the pursuit of those values undermines their happiness or their pursuit of other values even more important to them. In such cases, we say an individual has undermined *herself*. So a commitment to looking at how a decision will impact a person's welfare over time does not commit us to saying that any particular value is bad. We are seeking to identify a kind of inconsistency—it is just that here we are looking at an inconsistency that plays itself out *over time*. It is an inconsistency between some of a person's values or desires (those she would act on now if allowed to) and the rest of her values and her overall welfare.

Still, it may seem that such a direct assessment of welfare must be paternalistic in an objectionable sense. However, although I grant that requirement (1) is not ethically sound on its own, it is not yet clear that we can label the problem it poses “paternalism.” Recall, that as most theorists use the term, not all paternalistic actions (actions taken against a person’s will for that person’s benefit) are bad. 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 directed at a competent adult. That won’t help us here, since who is competent is the issue we are struggling with. Labeling the practice “paternalism” at this point begs important questions.

However, there *is* a moral problem here. It is not a problem of judging other people’s values, nor is it yet clear whether it is an instance of morally objectionable paternalism. However, the real moral issue with (1) by itself becomes clearer if we stop to think about the primary value that paternalism offends against, namely, free choice in the self-regarding sphere.

Recall that the freedom we grant to competent adults is the freedom not only to choose well, but also to choose badly. It includes the freedom to make prudential mistakes, and many people actually make them. Recall as well, that one constraint on any theory of capacity is that most ordinary adult human beings must count as competent most of the time. This suggests the following: As a society we have agreed that ordinary adults have a degree of decision-making capacity that, while not perfect, is *good enough*. But if this good enough ability to look out for oneself warrants the label “competence,” then an individual with demonstrably less of an ability to look out for herself than the ordinary adult should count as incompetent. In short, I am suggesting that it would be okay—not objectionably paternalistic—to intervene with an individual’s choice if, in addition to thinking that she was making a serious prudential mistake, we *also* had good reason to think this individual *was more likely to make prudential mistakes than an ordinary adult*. But notice that this just is requirement (2). In short, without (2) requirement (1) sets the bar too high,

allowing us to intervene whenever we have good evidence that someone is making a serious prudential mistake. That's too much intervention. But in conjunction with (2), (1) it does not have this feature.

The claim that we should factor into assessments of capacity how a choice will affect the person's welfare is not new. However, my suggestion that we give welfare this particular circumscribed role is. And it is a difference that makes all the difference. Recently Pickering et al. (2022) argued that "in some cases [a person should] be judged incapable of making [a] decision *because of the harmfulness of the decision.*" Noting the common distinction between process and outcome and the assumption that we must only assess process, they argue that in some cases the outcome of the decision *must* be considered if we are to arrive at proper assessment of capacity. Up to a point, I agree. But by itself, their suggestion is problematic for the same reasons my requirement (1) would be problematic on its own. We need to limit how often outcomes are allowed to play a role, and we need a justification explaining why the limit is set the way it is. We do not want to turn the freedom to decide for oneself into the freedom to decide *as long as one doesn't choose badly.* My two principles working together ensure that the fact that someone is making a serious prudential mistake only carries weight in conjunction with (2).

Requirement (2) appeals to the idea of a "condition" known to increase the likelihood of prudential mistakes. To apply (2) we would need to acquire evidence over time about the prevalence of serious prudential mistakes in certain populations. While not always possible, this might be possible in certain cases. For example, consider again anorexia nervosa. Refusal of treatment, including refusal of life-sustaining treatment, is relatively common among this population. It therefore seems likely that we could establish the following: choices resulting in preventable death are more common among people with anorexia than among the general population. This is not enough, however. We also need support for the claim that a significant number of these deaths are serious prudential mistakes as defined here.

When it comes to mental disorder, however, our sense of whether a choice is a mistake may

depend on how likely full recovery would be *if* the person lived. Serious prudential mistake occurs when someone rejects a life that, were they to live it, would subjectively offer *them* more good than bad, and it is easier to see that a choice is of this sort when we can rule out the possibility that the life will be plagued by the suffering characteristic of much chronic anorexia. Given that chances of full recovery from anorexia are highest (and reasonably high) for those in the first few years of illness (Treasure et. al 2015), this suggests it may make most sense to focus on a sub-set of patients with anorexia. Assume then that the relevant population is not all people with anorexia nervosa, but people who are at most a few years from diagnosis. Among this group it is plausible that the higher rate of choices resulting in death *is* a higher rate of serious prudential mistake. Further support for such a conclusion might come from additional research, e.g. following those who recover fully and studying their quality of life. Assuming the general quality of life in this population is high (or at least not low), this would further support the idea that in throwing away their lives, those who could have recovered but refused treatment made a poor choice.

As with requirement (1), there are ethical reasons why we should not appeal to (2) by itself. We saw earlier why it is important not to base a judgment about competence on a diagnosis alone. That would restrict too many choices of too many people for no good reason. When paired with requirement (1), however, requirement (2) does not do that. The requirement that the individual have a condition known to increase the likelihood of serious prudential mistake only becomes operative *if in addition* there is good reason to think this individual is making a serious prudential mistake *right now*. Finally, an advantage of this approach is that, because (1) focuses on *serious* prudential mistakes, many other less consequential decisions remain completely untouched by the framework, ensuring that the subject remains free to decide these matters as she sees fit.

Work would need to be done to define more precisely what a “condition” is for these purposes. I certainly intend it to cover a range of mental disorders, particularly given that many of the problem cases for the current framework arise among patients with such disorders. However, whether or not someone like Donna counts as having a mental *disorder*, it seems that she is in a

recognizable kind of situation—the immediate aftermath of receiving devastating news— typically thought to induce a temporary period of poor decision-making. However, I am not suggesting we rely on what is typically thought. Presumably decision-making in such circumstances has been studied and could be studied more to acquire the kind of evidence needed to support this.

§5.0 Conclusion

Various “conditions” involving mental disruption lead people to make seriously poor, self-harming decisions even from their own subjective perspective. Yet the dominant framework—the four abilities model—cannot capture all of these cases. 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*. The familiar four abilities are necessary but not always sufficient for establishing whether someone has capacity in this sense. I have proposed two new requirements which, if used together—*both* must be satisfied—can allow us to recognize cases where capacity is still lacking despite the presence of the minimal four-abilities. Requirement number (1) states that there must be good evidence for thinking that the patient is making a serious prudential mistake here and now. By itself it is ethically problematic, because it sets the bar too high. Requirement (2) states that the patient must be known to have a condition that, in turn, is known to make those who have it more likely to make prudential mistakes than ordinary people. By itself it is ethically problematic because it violates diagnostic neutrality. Together, however, the requirements offer a perfectly ethical addition to the four abilities framework—one that could greatly improve capacity assessment in certain kinds of difficult cases. For this to happen, however, we would need to develop guidelines for the interpretation of the requirements, collect data about outcomes, and develop various safeguards. I hope that enough readers see the promise in the approach to think that such work is worth undertaking.

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.
- 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., 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.

- 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.