

Clarifying Capacity: value and reasons

Jules Holroyd

**[this is a pre-publication copy, please do not cite without permission. Forthcoming in
Autonomy and Mental Health, ed. Radoilska, L. (anticipated 2011)]**

The aim of this paper is to clarify what is involved in the notion of capacity as used in the Mental Capacity Act of 2005. This act (hereafter MCA) sets out the conditions for ascertaining whether an individual lacks the mental capacity to make a decision. Where there is evidence for an individual lacking capacity, and further inquiry confirms this to be the case, then another individual is assigned to decide in the individual's best interests. Because judgements about whether someone has or lacks capacity can have significant consequences, in particular in the context of healthcare, it is crucial to come to a full and accurate understanding of what it is to have - or lack - capacity. Considered in the context of mental health care, we will see that ascertaining whether an individual has capacity is not only an ethical matter concerning the avoidance of harms, but also has political implications concerning what choices or conceptions of value the state should via the health service, permit.

I will be asking whether meeting the conditions set out in the MCA requires certain evaluative commitments. This question is particularly testing in the context of issues that arise concerning mental health, where what is believed relevant or given weight in making a decision appears to be bound up with the mental health problems about which decisions are being made. In this paper I will first clarify and elaborate on the claim that the conditions for capacity as set out in the MCA are value-laden (cf. Owens et al, Freyenhagen, Richardson, and Hotopf 2009) (section 1-3). Then I will show that whilst the conditions are indeed value-laden, and presuppose that certain evaluative commitments are held by capacitous individuals, significant difficulties arise in attempts to rationalise value-laden judgements about capacity, and much work still remains in this regard.

How is capacity relevant to the various understandings of autonomy? One might view capacity (with respect to a particular decision) as coextensive with being autonomous (with respect to that decision) - if so, then given the role that the notion of capacity is playing, it is clear that the particular target concept of autonomy at work is concerned with the boundaries of interference and paternalism. If one does not take capacity to be coextensive with autonomy, then individuals might be autonomous whilst lacking capacity, or vice-versa (depending on the notion of autonomy adopted). Discussion of which strategy one might adopt is deferred for another time.¹

1. Mental Capacity

Whilst it is usually appropriate for adults to make decisions about what kinds of medical treatments they undergo, sometimes impairments are suffered - either temporary or permanent - which render an individual unable to make such decisions. The Mental Capacity Act 2005 sets out the conditions under which it is appropriate to regard an individual as lacking the capacity to make a particular decision about treatment. It is important to note that the MCA is different from the Mental Health Act (2007), which sets out the conditions under which patients with mental illnesses may be detained for their own health or safety, or that of others, for the provision of treatment for the mental disorder. The two acts serve quite different purposes; individuals who have not been diagnosed with a mental illness may fail to meet the conditions for capacity - if they are in a state of temporary confusion or debilitation or are in a coma, say. Likewise individuals who have been diagnosed with mental health problems may still meet the conditions for capacity. However, as we will see, suffering from certain mental health problems or illnesses can make meeting these conditions difficult; patients who suffer from dementia may be unable to retain information in the way required; patients suffering from delusions may be unable to understand the relevant information, for example. I will focus in later sections on sufferers of anorexia nervosa, and what we might say about the ability to weigh the relevant information in coming to a decision.

¹ See also the contributions to this volume by Jane Heal and Hallvard Lillehammer.

As stated, the MCA specifies that an individual lacks the capacity to make a decision if:

'at the time he is unable to make a decision for himself in relation to the matter because of an impairment of, or a disturbance in, the functioning of the mind or brain.' (MCA, 2.(1))

An individual is in turn judged unable to make a decision about such a matter, if he is unable

- a) 'to understand the information relevant to the decision
- b) to retain that information [for sufficiently long to make the decision, at least]
- c) to use or weigh that information as part of the process of making the decision, or
- d) to communicate his decision (whether by talking, using sign language or any other means)' (MCA, 3. (1)).

One of the striking principles stated in the MCA is that: 'a person is not to be treated as unable to make a decision merely because he makes an unwise decision' (1.(4)). In so claiming we are encouraged to consider these conditions as content-neutral, or procedural - whether an individual has capacity should not depend upon the content of her decisions, or commitments.

However, in a recent discussion of the MCA, the authors conclude that 'if psychiatry aims at a completely value-neutral or even value-free conception of mental capacity it will come unstuck' (Owens et al, 2009, 100). If these authors are right, then it is important to work out precisely at what point values play a role in the conditions for capacity. Ascertaining this is important for three reasons: first, to clarify what considerations are relevant to judging capacity, especially in difficult cases - for some such cases seem to hinge upon whether the individual endorses certain values; second, understanding the relationship between autonomy and capacity requires first being clear on what the conditions for capacity demand; third, clarity on the evaluative content of the conditions

may have implications more broadly for public health ethics.

I will be primarily concerned with the first of these considerations, leaving the other two for detailed discussion elsewhere. It is clear that it is of crucial importance that the conditions for identifying when an individual lacks capacity are properly understood and applied; misapplication of those principles will lead to medical paternalism, infliction of moral harm, and imposition of the (risk of) physical harm attendant upon most medical procedures. Whether or not the conditions are value-laden also intersects with concerns in public health ethics. When it comes to which choices to respect, many have argued that the state should not make such decisions based on whether those choices cohere with certain state-sanctioned values; that individuals should be free to pursue whatever conception of the good they choose, so long as it does not harm or wrong others.² It will be important, then, to ascertain whether the conditions for capacity are value-laden, and in what way: might being judged to have capacity ultimately turn on whether one accepts certain values?

2. Evaluative components of MCA

An agent is deemed to lack capacity insofar as she is unable to do any one of the four things set out in the MCA. Precisely what is involved in meeting these conditions needs considerable unpacking. I aim to focus on the extent to which being able to understand information relevant to the decision, and to weigh that information in making the decision, are value-laden. Do judgements about when an individual is able to understand or weigh information in deliberation depend upon which values she endorses? This focus coheres with some aspects of Owens et al.'s treatment of the MCA. The authors identify four dimensions in which the conditions for capacity rely on value-laden, rather than purely descriptive criteria. These include:

² For a discussion of when and whether it is appropriate to override individuals' choices, see Dworkin (1983, 1972), Christman (2004), Oshana (1998). The target of autonomy in much of this writing is that which sets the boundaries for paternalism. See Arpaly (2003) for a taxonomy of the different notions of autonomy. For discussion of whether the state should be in the business of promoting certain values in healthcare, see Radoilska (2009).

- i. The patient's having 'insight into illness'
- ii. The patient's evaluations and the role these play in their reasoning
- iii. The emotional states of the patient and the role these play in practical reasoning and decision-making
- iv. The gravity of the decision, and the risk of falsely judging capacity when the potential harms are great. (2009, 92, 95-100)

The first two of these dimensions, the patient's insight into illness and her evaluative commitments, correspond to the two aspects of the MCA I have picked out as warranting further scrutiny; the understanding and weighing requirements. I will focus on the first two aspects, and set aside the other two, although these also deserve further scrutiny.³

2.1 Understanding relevant information

The text of the MCA adds that the information relevant to the decision includes (but as we shall see is not limited to):

³ A few notes on the latter two conditions are in order: first, the authors focus on the impact of 'altered emotional states' upon the individual's decision making process. Such emotional states may influence 'in a detectable and identifiable way, the meaning, value and weight given to treatment risks and benefits, such that the patient may be unable to appreciate the benefits side of the equation, or may become unduly concerned about the risks.' (2009, 98) The question of what the appropriate emotional responses to certain benefits or risks, and what amounts to undue concern, becomes salient. Whilst some of these issues will be peripherally addressed in the discussion of the patient's evaluations, the matter of the appropriateness of certain (strengths of) emotional responses and their relevance to capacity will not be tackled head on here. The question of whether any the agent must hold certain evaluative commitments (believe certain things to be valuable) in order to have capacity can be answered without settling this matter, and insofar as emotional responses are attached to certain evaluative dispositions (and perhaps are constitutive of them) we will touch upon this matter indirectly. Moreover, ascertaining whether capacity requires certain evaluative commitments may then help with later assessments of the rationality of certain emotional responses to the objects of those commitments.

The fourth dimension picked out by Owens et al focuses on the evaluations of the assessors - the weight assigned to the risks and benefits of the decision undertaken, and how this may affect where the bar for capacity is set. They write:

'to judge incapacity falsely results in treatment in the patient's best interests, whereas to judge capacity falsely may result in serious harm or death which may have been preventable and which the patient did not autonomously choose. It is more valuable, by implication, for risky decisions which are incapable to be blocked than risky decisions which are capable to be permitted. This makes it imperative, so the argument goes, for assessors of capacity to be surer about capacity in the context of a risky decision.' (2009, 99-100)

Note that here, the claim is that despite the moral and physical harms that may ensue, when the stakes are high a false negative judgement of capacity is better ('more valuable') than a false positive. Once again, this issue is not wholly isolated from the matter of what evaluative commitments a patient must have in order to meet the conditions for capacity, for her evaluation of the risks may differ from that of the assessors. Our focus for now is not primarily upon the evaluations of those assessing capacity, but rather on what evaluative commitments an agent may be required to have if she is to meet the conditions for capacity.

'information about the reasonably foreseeable consequences of a) deciding one way or another, or b) failing to make the decision'. (MCA, 3.(4)).

What more is required for understanding the information relevant to a decision? In this section I will argue that the understanding condition requires that individuals subscribe to a certain range of values, and that this is most clearly the case in relation to the requirement for 'insight', namely, understanding that one is ill.

What is deemed to be relevant to the decision in hand will, in part, be determined by the evaluative commitments of the agent. For example, that one course of treatment violates a religious doctrine (to draw on an example to which we will return) might be considered relevant by a patient who is committed to that religious system and its doctrines, but not deemed (independently) important by some medical practitioners. The grounds for regarding such information as relevant (and the patient's attendant beliefs that receiving certain treatment would be sinful, say) is not independent of the evaluative weight that the patient gives to different considerations; the considerations to which an individual accords significant weight will presumably correlate with what she regards as relevant. I return to the issue of weighing the information in section 3. The considerations that arise in this regard will not only be relevant to decisions made in medical contexts, but will pertain to judgements about capacity in relation to decisions in other domains also. In the context of healthcare - particularly mental healthcare - some specific propositions are regarded as relevant, and this section will focus on the extent to which these are value-laden.

2.2 Insight into illness

Owens et al. identify 'insight into illness' as a value-laden dimension of capacity (2009, 95). The notion of 'having insight into illness' pertains to whether or not the patient recognises that she has

an illness. Clearly, such a recognition is necessary in order to make sense of why certain treatment options (or their refusal) are on the table – that is, why any decision needs to be made at all. Unless one understands that one is unwell and that treatment may remedy this, the only sensible option will be to refuse treatment. Thus we see that having insight falls under the specification of the MCA which requires that one ‘understand the information relevant to the decision’. What is required in having insight?

Owens et al refer to ‘a judgement about the patient’s ability to recognise certain experiential states as pathological. And such judgements are value-laden’ (2009, 95). The focus of their discussion is on the evaluative nature of the judgements (by practitioners) about whether an individual has insight. We should want to know a great deal more about the ways in which such a judgement is value-laden. I will focus on the question of whether an individual’s having an awareness of their illness requires that they accept certain evaluative propositions or values more broadly (in particular, those which correspond to the practitioners judgement regarding what it is to be in good health, or ill). Whilst this is important for the cases that Owens et al. raise - cases in which a feature of the illness is that the symptoms are not recognised as such (2009, p.95) - we will see that the evaluative requirements are necessary even in run of the mill acknowledgements of illness.

A first gloss on what ‘insight’ requires is as follows: a patient suffering from an illness with symptom’s S ought, insofar as she has insight, to subscribe to the beliefs:

- (a) I am experiencing (symptom) S, and
- (b) S is indicative of a disease/illness

However, articulated in this way, it is not clear precisely what the evaluative element is: (a) and (b) may appear to be purely descriptive statements; one about the patient’s experiential states, the other

about the relationship between these states and disease. If this is so, then it is not clear that insight is an evaluative matter, but rather a matter of simply forming certain beliefs about one's own states and about what these states mean in terms of one's state as diseased or not.

However, in such judgements there are in fact (at least) two evaluative aspects. Fulford (1989) has argued for an understanding of the notions of health and disease as evaluative. Roughly put, on his view to 'be in good health' or 'to be well' is to meet certain standards of functioning designated as valuable states. To suffer from illness is to fall short of these evaluative standards (first evaluative component) in a way that is disvalued or regarded as undesirable (second evaluative component). If this is right, then having insight into illness will involve recognising:

- i) I am experiencing (symptom) S, and
- ii) S indicates a negative departure from the norms of good health, H.

If this is right, then having insight into illness (and understanding this bit of relevant information) requires accepting certain normative ideals of health.

2.3 The evaluative content of assessments of health

Does this mean that coming to an understanding of the relevant information is only possible insofar as one subscribes to certain ideals of health? If so an individual's disagreement over a state being pathological would mean that she lacked insight, and thereby lacking the capacity to make the relevant decision. But Fulford's understanding of illness as evaluative acknowledges that there might be some variation in the relevant ideals and evaluations. Whilst there might be general agreement on the existence of a certain experiential state, whether that state gives cause for diagnosis of an illness or disease will sometimes depend upon the patient's evaluation of it, as Fulford, Dickenson and Murray write:

what is a problem for one person – having an ache or pain, being a certain weight, having a given level of energy or a particular sleep pattern – may not be a problem for another person ... medical diagnosis is ... a matter of negative evaluation (2002, 6)

On this view, being ill will involve the patient taking a negative evaluation towards her experiential state. An agent might evaluate a state negatively because it departs from certain norms of health that she accepts (and that others do not). Alternatively, she might note that it departs from these norms, but not find that departure problematic. Thus two individuals with similar experiential states may differ with respect to whether they regard themselves (or are diagnosed as) suffering from some illness or defect of health. The differing conceptions of the norms of health might lead A and B to differ in their judgements about whether (e.g.) an erratic sleep pattern is a health problem. Or such disagreement might be characterised differently: even if A and B accept the same norms of health, B might not find the departure from the norms she accepts as problematic.

If there is sometimes room for divergent understandings of the norms of health, or of evaluations of departures from them, this is relevant to the evaluative nature of the condition of understanding 'insight into illness'. It may be that an individual does not accept a certain experiential state is symptomatic of an illness. This may be because that individual is unable to recognise that symptom as a departure from the norms of health she accepts, and so cannot understand that she is ill (hence cannot meet the 'understanding' condition of the MCA).

However, it might be that an individual's conception of the norms of health differs from those norms accepted by many others (it might also be that the importance she attributes to some state deemed by some pathological differs significantly - a point to which I return in section 3). As Fulford et al. note, many norms are widely shared: a heart attack is generally accepted as bad for a person (2002, 8). Plausibly, there are some such states which are bad for a person irrespective of their evaluation

of it. But it will not be the case that all norms of good functioning are shared, nor that all symptoms or experiential states will be bad for a person independently of their evaluation of it.⁴ There is scope for divergent conceptions of some norms, as well as divergent evaluations of departures from the set of norms that agents accept.⁵ This value laden understanding is supported by Derek Bolton's claim that, in the context of mental health care:

a primary task should be to define what the problem is and for whom, according to whose values, in an active collaboration with the individual or others concerned (2000, 151).

Thus understood, it appears that there are a range of ideals that might be accepted consistently with having mental capacity: people diverge in their ideas about what good health involves (even if there are some norms on which all (should) agree). Moreover, that individuals do diverge in their views of good health, and that having such preferences ignored can be upsetting and distressing (see Lillehammer, this volume), provides pragmatic reason in favour of operating with a more pluralist understanding of the norms of health.

Plausibly, then, for an individual to have insight, she must make evaluative judgements concerning the norms of health to which she subscribes; and concerning whether her experiential states are divergences from these.⁶ Whilst value-laden, then, we haven't yet seen that making judgements about capacity would be determined by whether or not an individual accepts certain values or not; rather, it is a matter of looking for a coherent understanding of how an individual's experiences fit

⁴ It is worth also noting the distinction, discussed by Fulford (2001), between dysfunction and illness; the text books may set out what scientists, with complex understanding of physiology, identify as good or bad functioning. Whether or not a divergence from this function is a divergence from the norms of health that an individual accepts, or evaluates negatively, will determine whether or not a medical diagnosis of illness is appropriate.

⁵ Strictly speaking, we should understand the patient's evaluation as in accordance (or not) with the norms of health H that she would accept, were she to reflect upon them under optimal conditions. This is required so that the analysis is not hostage to the patient's current mental states (which may in fact be subject to distortion or mistake). This understanding is different to, but (I believe) consistent with, the account of insight that Fulford develops – he understands lacking insight in terms of misconstrued attributions – for example, seeing some trait as done to oneself, rather than as done by oneself (or vice versa). See Fulford (2004) esp p.56. Such wrongful attributions would presumably not be recognised as departures from the endorsed norms of health.

⁶ As Fulford (2004) notes, some failures of insight will be due to the failure to recognise an experiential state as a divergence from endorsed norms of health, but rather to attribute it to e.g. one's environment (to which one's response is normal).

with her conception of health.

But can any conception of health be endorsed?⁷ I've noted that there are strong intuitions that some norms of health are just wrong. Further intuitions can be found in various feminists' arguments for women's rights in relation to health: Chambers considers women who, having undergone female genital cutting rituals suffered significant health complications but did not regard their bodily functions (prolonged menstruation and urination, painful penetration and childbirth) as abnormally difficult (2007, 213). The intuition is that it is simply mistaken to fail to regard such difficulties as a departure from good health (rather than a different but reasonable conception of health). Such an intuition is vindicated by the fact that, with more information, such women changed their views: 'once women do realise that they have been harmed by FGM, they are keen to put in place the village-wide declarations that are necessary to abandon the practice, as fieldwork demonstrates' (Chambers 2007, 214). Whilst on the one hand consideration of such cases appears to support the thought that there are limits to the norms of health that can be accepted, one might maintain that such a conclusion does not rely on assumptions about what is and what is not a norm or ideal of good health. (Doing so would yield a value-laden notion of capacity, as whether an individual has insight will depend upon whether she accepts certain ideals of health, and is able to recognise her symptoms as departures from these.) Rather, content-neutral conditions pertaining to the information an individual has available to her whilst formulating understandings of good health can be appealed to: the women's understanding was mistaken, in being formed without relevant information (rather than mistaken, because it was substantively wrong).⁸ Perhaps this kind of content-neutral condition can account for intuitions about which norms of health are mistaken or wrong. However, when we look to work in mental health care, it appears that there are some cases in which, even when equipped with all the relevant information, seemingly problematic norms of

⁷ Parallel concerns may arise in other domains in which considerations of capacity arise (in social care, or legal settings, say, where decisions are to be made about responsibility for welfare, or property): can any conception of the good life be consistent with capacity, or do some commitments indicate a lack of understanding of the relevant information? Thanks to Edward Harcourt for pressing this point.

⁸ Compare some of Dworkin's (1988) necessary conditions for autonomy, requiring that endorsements are formed with relevant information, and under conditions of procedural independence.

health are subscribed to.⁹

Some sufferers of anorexia nervosa reject the ideal of a certain weight that affords good functioning (physically and mentally). Certain experiential states may be recognised (say, dizziness, weakness, resulting from malnourishment), but not regarded as a departure from the individuals' accepted norms of health (which do not feature adequate nutrition). Is this indicative of lack of capacity to decide about treatment? This brings us to the weighing condition.

3. Weighing information

In addition to understanding the relevant information – understanding, say, that one is ill, or that according to certain standards of wellness one's experiential states are divergent from these norms – one must also be able:

- c) To use or weigh that information as part of the process of making the decision (MCA, (1))

What might we look to in order to determine whether an individual is able to weigh information? The ability to weigh information as part of a deliberation process is in the first instance subject to certain formal constraints - such as consistency and transitivity. Inconsistent evaluations thwart stable processes of deliberation and weighing, as does the inconsistent ranking of options in order of value. In the instance in which an agent violates these constraints, we may have reason to suppose that they are unable to weigh the information in a manner required for capacity (or perhaps may be unable to retain information long enough (as required by the second condition) in order to

⁹ A further concern is that conditions pertaining to having the relevant information may also be practically difficult to implement. Consider a condition pertaining to whether or not one had the relevant information in forming a norm of health. Most individuals do not have reflective access to when they formed views on what good health consists in, less still what influenced them in doing so; nor is such information likely to be readily available from others. The point is not that this condition may not ultimately be able to identify circumstances that threaten the capacity to make decisions (perhaps undermining insight or the ability to weigh information). Rather, the worry is that the application of those conditions to ascertain whether an individual has capacity or not will not be practically possible: it will require information about processes according to which individuals formed desires or endorsed values - information that is not easily discerned, least of all in contexts where decisions may be required with some urgency.

make consistent evaluative rankings). Certainly, failure to meet such constraints would lead to erratic patterns of valuing, which would be difficult to reconcile into a stable decision about treatment options.

Such constraints, however, do not require that any *particular* weights are assigned to certain options or experiential states: these formal constraints are consistent with assigning a low value to some physiological functions, and a high value to treatment avoidance (insofar as such commitments are consistent with other values held). Does weighing information require that certain specific evaluative commitments and rankings are held? For example, would refusing treatment because one cannot see any value in continuing what would be an otherwise physically healthy life indicate that one is not properly weighing the information, giving too little weight to the harms of leaving the condition untreated? Or would refusing treatment for severe malnourishment due to anorexia nervosa in order to maintain a low weight and avoid food consumption indicate that one is improperly valuing nourishment and giving too much weight to food avoidance and maintaining a low weight? Would refusing life-saving treatment on the basis of religious doctrine? I want to focus in this section on the case of anorexia nervosa as it locates our discussion in the context of mental health, and provides a particular challenge for content-neutral understandings of capacity.

Tan and Hope (2008)'s research implies that certain evaluative weightings are required for capacity, and that this can be seen in considering the competences and capacities of patients with anorexia nervosa. Such patients are described as (frequently) meeting the formal conditions for capacity, manifesting high standards of rationality. However, these individuals, it is argued, 'overvalue' maintaining a low weight, leading them to assign to it a greater evaluative weight than that assigned to participating in treatment programmes that would ensure a healthier body weight. In an earlier piece (Tan, Hope and Stewart 2003) the formal competence of the patients is described; however, it is remarked that anorexic individuals experience difficulties in their substantive valuing:

all the participants were already highly conversant with the facts of their disorder, the exercise of going through information about anorexia nervosa and its treatment [...] was experienced as onerous and patronising to the participants and awkward and painful to carry out for the interviewer. This suggests that the standard concept of capacity to consent to treatment, as being one of understanding and reasoning [...] may not be relevant to the difficulties that these participants may experience in their decision making.

One important area, which emerged from the qualitative analysis, is attitudes to death and disability. Treatment refusal may occur, not because the patient wishes to die, but because of the relative unimportance of death and disability as compared to anorexia nervosa, or because of the particular meaning death and disability may acquire in the context of anorexia nervosa (2003, 704).¹⁰

In remarking on the ‘relative unimportance’ to these patients of death and disability, Tan et al. identify the patterns of valuing which diverge from those deemed normal. Many people endorse, as a norm of health the maintenance of a body weight that permits daily activities and does not risk disability or death. But such a norm does not figure - or not significantly - in the norms accepted by the patients described. We tend to think that someone who weighs being thin or avoiding food over death or disability is not just endorsing a different set of values; rather, they are making an evaluative mistake. This seems to indicate, then, that our understanding of what it is to be able to weigh information in a way relevant to capacity concerns the assignment of *appropriate* weights to certain objects or options or norms, which determines how one should rank them: to value life over food avoidance; to value treatment over risk of death or disability. Similarly, it has been argued that

¹⁰ Some anorexia sufferers may have false beliefs about their size or weight, or experience significant emotional resistance to eating or weight gain. In such cases, they may not meet the formal requirements; or it may be that conditions that make reference to emotional states (as recommended by Owens et al.) are relevant to capacity. My claims in this section, then, are of most relevance in the cases described by Tan et al., where formal conditions are met, and insofar as the conditions for capacity do not make explicit reference to emotional states. I thank Agnieska Jaworska and Lisa Bortolotti for raising these considerations.

depression involves undervaluing of continued existence.¹¹ And controversial cases in which individuals with religious commitments refuse potentially life-saving treatment centre on whether the weight given to considerations of faith and religious doctrine undermines ability to properly weigh other relevant information (cf. Martin, 2007). If this is so, then meeting the conditions for capacity seems to require not merely that the agent hold some evaluative commitments (as is required for understanding the information relevant to the decision); but also that the agent hold some specific and substantive evaluative commitments.

I take it there are two key thoughts underpinning these judgements: first, that an agent who makes such divergent evaluations is not merely engaging in different patterns of valuing, but rather in distorted patterns; she is not simply diverging from the (statistical) norm, but is somehow getting something wrong. Again, an evaluative conception of this condition for capacity, which rules out at least some value-commitments, must be endorsed if this thought can be accepted. The second thought is that it might be appropriate to make decisions on behalf of an individual, if her distorted understanding means she is unable to weigh information in making a decision (and where the consequences of that decision are grave).

3.3 Consequences of value-laden conditions for capacity

I haven't offered a watertight argument for the conclusion that meeting the weighing condition for capacity requires certain evaluative commitments; rather, I have shown that the structure of the understanding and weighing conditions presupposes that certain values are held, drawn out strong intuitions in favour of this conclusion, and appealed to empirical research findings. There are strong reasons for supposing that both the understanding and the weighing conditions require certain evaluative commitments: insight into illness requires that some values are endorsed; the weighing condition appears to require that this endorsement is of values, given appropriate relative weight, from a limited set. It is plausible that those values in play might be plural, leaving some room for

¹¹ Moorhead S, Turkington D. Letter: Role of emotional capacity in consent should be clarified. *BMJ* 2002;325:1039

individual divergences regarding what is of value. But certain beliefs or commitments, or rankings of them, will be ruled out.

This preliminary conclusion may be an uncomfortable one: it is certainly in tension with the stated principle that judgements of capacity should not be determined by the content of an individual's decision. The following consequences of a value-laden set of conditions stand out:

i) If over-valuing some option is taken as evidence that an individual cannot weigh the information in coming to a decision, some 'unwise' decisions will be relevant to ascertaining whether an individual meets the conditions for capacity. Such a conclusion implies that whether an individual's values are unorthodox may be relevant to her decisional capacity.

ii) This also suggests that there may be an asymmetry in ascertaining capacity: individuals who make 'orthodox' weightings of value and relevant information may not be pressed to explain or justify their decisions in a way that an individual who is regarded as 'over' or 'under' valuing some considerations.

iii) Further, this conclusion poses a challenge for those who defend a value-neutral understanding of the state's obligations in public health care; we will need to see how such a framework can be consistent with the MCA, which seems to rely on the conceptions of value an individual endorses in making assessments of decisional capacity.

iv) Finally, if the MCA does contain value-laden criteria in this way, then further discussion will be required about which values are consistent with capacity and which are not. The alternative would be to re-formulate the MCA so as to avoid reliance on a 'weighing condition'.

However as we will see in the following section, there is significant difficulty in systematizing intuitions about what undermines decisional capacity.

4. Systematizing intuitions

The preliminary conclusion is that, as Owens et al. argue, it is mistaken to maintain that the conditions for capacity are value-free. I have drawn out two ways in which the conditions for capacity require that individuals accept certain values, most notably the weighing condition. If it is right that assessing decisional capacity depends upon an individuals' evaluative commitments in this way, then it is essential that some systematic way of ascertaining which value commitments are consistent with being able to weigh information, and which threaten to undermine this ability.

As it stands, intuitions seem to pull in different directions: it appears intuitively plausible that over-valuing food avoidance or under-valuing continued existence thwarts the ability to weigh information relevant to treatment decisions. On the other hand it is less intuitively compelling to think that under-valuing the risk of death or disability due to a commitment to religious doctrine undermines decisional capacity (although anecdotally, intuitions seem to vary significantly on this).

In a recent Hastings Center Report, Adrienne Martin (2007) has offered arguments which may account for such conflicting intuitions. In this section, I address these claims, arguing that they cannot explain our divergent intuitions. These arguments are also important in addressing which considerations, beyond decisional capacity, may speak in favour of respecting treatment choice.

If one does not share the intuition that religious believers lack decisional capacity, and rather should have their choices respected, this would speak against a value-laden understanding of the conditions for capacity: over-valuing other options relative to avoiding death or disability cannot be what

undermines capacity. The treatment decisions of certain religious individuals involve certain weighting or assignments of value that may, to non-religious individuals, be as distorted and misguided as those evaluations of the anorexic patient. For example, in cases in which life is threatened unless a blood transfusion occurs, some followers of the Jehovah's Witness faith will assign a greater value to maintaining a pure soul (believing an interpretation of the bible according to which God prohibits the transference of blood) than to remaining alive.¹² For those who give no weight to the commands of God or the purity of the soul, such evaluations are misguided and distorted.

However, the tendency to see these beliefs and evaluations as undermining of capacity is weaker than in the case of the anorexic patient; whilst one might disagree with the evaluative weighting, one might nonetheless maintain that, because of her religious commitments, it is appropriate to respect this divergent evaluation. Such cases might be appealed to in support of the claim, then, that evaluative commitments do not inform judgements of capacity - otherwise intuitions and decisions about some such religious believers would pull in the same direction as those regarding the anorexia sufferer.

Martin's arguments may provide a way of accounting for such divergent intuitions. She argues that even if individuals are lacking capacity, there may be other reasons for which it is important to respect their choices. These other considerations may be informing intuitions about the case of religious believers. If her arguments work, this would explain how a value-laden understanding of capacity can be consistent with divergent intuitions about which choices to respect. However I will argue that her arguments not successful, which leaves the situation somewhat vexed with respect to our intuitions about intervention and capacity, showing the need for more attention to which values are consistent with decisional capacity and why.

¹² It is important to note that many do not so believe, and campaign for the legitimacy of their pro-transfusion views within the Jehovah's Witness faith. See the website of the Associated Jehovah's Witnesses for Reform on Blood.

4.1 Respect for autonomy

In this paper, I have left open whether we should identify capacity with autonomy, or whether the two notions might come apart. The stance one takes on this will depend (in part) upon whether the conditions for capacity are value-laden (whether or not autonomy is value-laden or content-neutral is itself contentious). However, in order to base an argument for respecting the decision of an agent who lacks capacity in respect for autonomy, the notions of autonomy and capacity cannot be seen as coextensive. For if so, an agent who lacked capacity, would also lack autonomy; respect for her autonomy could not then provide an additional reason for respecting her choice.

Martin suggests that we consider the possibility that ‘even a person lacking capacity qualifies as autonomous’ (2007, 38), and that their status as autonomous might provide reason to respect a choice even if capacity is lacking (due to, say, ‘distorted’ evaluative judgements). Martin works with an understanding of autonomous agency as ‘coherence or consistency across one’s value hierarchy and how one acts in relation to that hierarchy’ (2007, 38) such that autonomous persons are unified in their endorsement of and action upon certain values (cf. Frankfurt, 1971). She suggests that certain evaluations based in false beliefs (about, e.g. valuing treatment refusal over acceptance, due to the belief that eternal damnation is a consequence of the latter) might nonetheless play an important unifying role for the agent. If this is so, then respect for that agent’s autonomy (her capacity for coherent evaluative endorsements which guide action) might mean respecting a choice that would not meet the conditions for capacity:

‘the requirement to respect autonomy may therefore require respecting some incapacitated decisions [...]. For some decisions, respect for autonomy might require that we respect her treatment decision based in her [false] belief, even when she ‘unreasonably’ retains that belief [...] – even when the belief renders her incapacitated’

(2007, 39).

Martin is here (with her reference to the unreasonable retention of belief rendering the agent incapacitated) supposing that some beliefs might get in the way of understanding the relevant information, or weighing it appropriately (she writes that the 'Jehovah's Witnesses religious beliefs prevent them from arriving at an even remotely accurate assessment of the risks of blood transfusion' (2007, 36)). But Martin claims that even if the agent does not meet the conditions for capacity – if her false belief prevents her from understanding or weighing the relevant information – it might nonetheless be appropriate to treat her as if she has capacity for the sake of respecting her autonomy.

Whilst this might be so, it will not help us in explaining any divergent intuitions about the anorexia sufferer and the religious believer; for we might suppose (plausibly, given what was said about the formal competence of many sufferers of anorexia) that the anorexic meets the relevant coherence and endorsement conditions in order to be considered autonomous, in this sense. An individual – particularly if she has suffered from anorexia over a prolonged period – might consider anorexia 'part of herself', identifying and endorsing the abstinent motives and evaluative commitments to thinness and food-avoidance, much in the same way that individuals regard other evaluative commitments as constituting their identity.

That anorexic individuals could meet these conditions on the structure of the will may place in doubt their sufficiency for self-governance or autonomy. Indeed, that anorexic patients might endorse at a higher order their desires for abstinence is precisely what leads Marilyn Friedman (1986) to express concern about such accounts: the first order desires for food from which the sufferer is alienated, she maintains, are ones we might be inclined to regard as her 'true' desires: authentic and constitutive of her 'true self'.

But on an account that privileges higher order endorsements, respect for autonomy might require respect for the anorexic's refusal of life-saving treatment, for the sake of her 'overvaluing' thinness or food avoidance. Insofar as this is counter-intuitive, it cannot be that these structures of endorsement (understood as autonomy) provide sufficient reason for respecting choice. For these structures of the will could be found in religious believers and anorexic individuals alike, and would provide reason for respecting the choice of both, even if both also lacked capacity.

4.2 Respect for institutions

An alternative argument Martin offers concerns whether an individual's 'distorted' evaluative commitments are located within established practices or institutions. She writes that:

‘it may be appropriate, even important, to evince attitudes of tolerance, admiration, or even respect for some of the communal practices people participate in, regardless of whether those practices or participation in them are connected to individuals' autonomy. Religion can be a deep source of meaning in individual and community lives' (2007, 39).

The thought here seems to be that respect for certain valuable practices requires respecting choices based on values embedded in those practices, even if the commitment to those values undermines the agent's ability to meet the conditions for capacity, or her autonomy. So supposing the believer's faith prevents her from meeting the conditions for capacity, there might nonetheless be reason to respect her choice for the sake of respecting or sustaining those practices that have informed her evaluations. This is because those practices perform other important functions, even if they also serve to undermine the capacity to make important decisions of those who partake in them.

This argument is problematic on a number of grounds. First, it expresses a worryingly instrumental attitude towards the individual whose decisional capacity is at issue; recall (as mentioned in section 1) that the decisions at issue in the medical realm may frequently be ones which involve significant risks and burdensome costs. To allow an individual who is not equipped to make such significant decisions to do so for the sake of preserving some traditions or practices is to use that individual as a mere means. This seems to be a significant failure of respect for that individual, and fails to demonstrate adequate concern for their welfare.

Second, the argument has strongly conservative tendencies; consider the practices described in section 2 of female genital cutting, which detrimentally impacted upon the health of those women who partook in those practices. Suppose (as is not implausible, but does not, of course, leave those practices impervious to critique) that these practices of female genital cutting serve some role in securing social cohesion, and have significance and meaning for those who participate in them: the practice is an initiation rite which secures for young women the status of eligibility for marriage, and confers some esteem upon them as well as those who perform the rituals.¹³ In the example we considered, many of the women had insufficient information to understand the impact of these practices upon their urinary and sexual functioning. With the relevant information, they revised their assessment of their experiential states in relation to a likewise revised conception of the norms of good health. It seems deeply troubling to maintain that, whilst lacking the understanding of the relevant facts to enable informed decisions, the value of those practices in maintaining social cohesion and individual meaning provides reason to respect a decision which supports that institution. The mere fact that there is a social practice, and that this social practice serves some useful function, cannot provide sufficient reason to respect choices that cohere with and reinforce those practices. This would be especially so when (as in Martin's example, and unlike the Sudanese

¹³ Indeed, in her discussion of practices of genital cutting in Senegal, Anne Phillips notes that the difficulties ending rituals of female genital cutting are essentially coordination problems: whilst no individuals were strongly attached to the practice, no community wanted their daughters to miss out on the social benefits accessible to those who underwent FGC. Hence the success of projects that signed communities up to a collective pledge to end FGC. See Phillips (2007), esp.46-47

women in Chambers' example) we have reason to believe that those choices are made by individuals who do not meet the conditions for capacity.

Finally, we can see the inadequacy of this argument by considering what might be said with regards the choices of anorexic patients. If the value provided by certain practices legitimises respecting choices irrespective of whether they are made with capacity, then if it were the case that the valuing of thinness were embedded in certain practices that gave significance and meaning to a community, then there would be reason to respect those choices. Indeed, one might argue that such communities and practices exist: perhaps with the existence of 'pro-ana' communities (where anorexic individuals provide support and motivation for each other in online forums, engaging in projects which have become known as 'thinspiration'), or in the increasingly prevalent social norms of thinness, reinforced daily on catwalks and billboards, and in the myriad women's magazines that provide normative frameworks in which many women find meaning. But to cite such practices as providing reason to respect choices that value thinness over life is as unappealing as citing practices of genital cutting as providing reason to respect choices that lead to damaging health consequences (although of course there are many significant differences between the cases).

In short, considerations that might be marshalled in favour of respecting choices based on divergent religious values do not seem able to provide sufficient reason for respecting those choices. This leaves things in a bit of a muddle: one explanation for the intuition that treatment which is in the best interests of the individual may be imposed on an anorexia sufferer who faces death or disability, it seemed, was that according to a value-laden conception of capacity, such an individual lacked decisional capacity, being unable to properly weigh information in coming to a decision. But nor, on such a value-laden conception, do religious believers, who privilege their commitments over life-saving treatment. But we (tend to) have different intuitions about whether such religious believers have capacity. These different intuitions cannot be explained away by appealing to the

value of respecting autonomy or institutions, as I have just argued. This is relevant to the MCA, because intuitions about the evaluative capacities of anorexic patients pushed us towards a value-laden conception of capacity, whilst intuitions about the religious believers seem to pull in the other direction. If it is accepted that the conditions of the MCA are value-laden, in the ways I have set out, then significant work remains in systematizing intuitions about which value commitments or rankings undermine decisional capacity, and which do not.

Conclusions

The considerations I have raised indicate that the conditions for capacity presuppose that individuals endorse certain evaluative commitments, both in regard to what information is relevant to the decision (in particular, relevant to insight into illness), and in regard to the ability to weigh information in deliberation. If this is right, then it is a mistake to suppose that whether or not one meets the conditions for capacity can be determined purely formally, without reference to the content of an individual's choice or commitments. I have drawn out certain consequences of this conclusion, regarding the different requirements that face individuals with conventional or unconventional value commitments; the consequences for value-neutral conceptions of public health; and the need to consider further which values are deemed to be inconsistent with decisional capacity. If such value-judgements do play a role in assessments of capacity, it is then important to open discussion of which values are to be respected and when, rather than to suppose that any choice is permitted, no matter how unwise or irrational. This is especially so as intuitions about which evaluative commitments hinder decisional capacity appear to conflict, and a value-laden notion of capacity seems to require that some of our intuitive judgements are revised or rejected.¹⁴

References

¹⁴ This paper has benefited greatly from discussions with Lubomira Radoilska and with the members of the reading group on Autonomy and Mental Health at the University of Cambridge, to whom I am grateful for their generous feedback. I am also grateful for feedback from Agnieszka Jaworska, Elizabeth Fistein, Fabian Freyenhagen and Lisa Bortolotti, and from the audiences at the conference on Autonomy and Mental Health at The University of Cambridge.

- Arpaly, Nomy. 2003. Varieties of Autonomy. In *Unprincipled Virtue*, Nomy Arpaly, 117-149. Oxford, Oxford University Press.
- Associated Jehovah's Witnesses for Reform on Blood. <http://www.ajwrb.org/about.shtml>
- Bolton, Derek. 2000. Continuing Commentary: Alternatives to Disorder. *Philosophy, Psychiatry, & Psychology* 7 (2): 141-153
- Chambers, Clare. 2007. *Sex, Culture, and Justice: The Limits of Choice*. Pennsylvania: Pennsylvania State University Press.
- Christman, John. 2004. Relational Autonomy, Liberal Individualism, and the Social constitution of Selves. *Philosophical Studies* 117: 143-164
- Christman, John. 1991. Autonomy and Personal History. *Canadian Journal of Philosophy* 21: 1-24
- Dworkin, Gerald. 1988 *The Theory and Practice of Autonomy*, Cambridge: Cambridge University Press.
- Dworkin, Gerald. 1983. Paternalism: Some Second Thoughts. In *Paternalism* ed. R. Sartorius, 105-112. Minneapolis: University of Minnesota Press.
- Dworkin, Gerald. 1972. Paternalism *Monist* 56: 64 – 8.
- Frankfurt, Harry. 1971. Freedom of the Will and the Concept of a Person. *The Journal of Philosophy* 68 (1): 5-20
- Friedman, Marilyn. 1986. Autonomy and the Split-Level Self. *Journal of Social Philosophy* 24: 19-35.
- Fulford, K.W.M. 2004. Completing Kraepelin's Psychopathology: Insight, Delusion and the Phenomenology of Illness. In *Insight and Psychosis*, ed. Xavier Amador, and Antony David, 47-66. Oxford: Oxford University Press.
- Fulford, K.W.M. 2001 "What is (mental) disease?": an open letter to Christopher Boorse. *Journal of Medical Ethics* 27: 80-85.
- Fulford, K.W.M. 1989. *Moral Theory and Medical Practice*. Cambridge: Cambridge University Press.

- Fulford, K.W.M., Donna Dickenson and Thomas H. Murray. 2002. Introduction: Many Voices, Human Values in Healthcare Ethics. In *Healthcare Ethics and Human Values*, eds. Fulford K.W.M., Donna Dickenson and Thomas H. Murray, 1-21. Oxford: Blackwell Publishers Ltd.
- Heal, J. Mental Disorder and the Values of 'Autonomy' *Autonomy and Mental Health* ed. L Radoilska, Oxford University Press.
- Lillehammer, H. Autonomy, Value and the First Person, *Autonomy and Mental Health* ed. L. Radoilska, Oxford University Press.
- Martin, Adrienne M. 2007. Tales Publicly Allowed: Competence, Capacity, and Religious Belief. *Hastings Center Report* 37 (1): 33-40.
- Moorhead S, Turkington D. Letter: Role of emotional capacity in consent should be clarified. *BMJ* 2002;325:1039
- Oshana, Marina. 1998. Personal Autonomy and Society. *The Journal of Social Philosophy* 29 (1): 81-102.
- Owens, Gareth .S., Fabian Freyenhagen, Genevra Richardson, and Matthew Hotopf. 2009. Mental Capacity and Decisional Autonomy: An Interdisciplinary Challenge. *Inquiry* 52 (1) 79-107.
- Phillips, Anne. 2007. *Multiculturalism without Culture*. New Jersey: Princeton University Press.
- Radoilska, Lubomira. 2009. Public Health Ethics and Liberalism. *Public Health Ethics* 2 (2): 135–145.
- Tan, Jacinta and Tony Hope. 2008. Treatment Refusal in Anorexia Nervosa : A Challenge to Current Concepts of Capacity. In *Empirical Ethics in Psychiatry* ed. Guy Widdershoven, 187-210. Oxford: Oxford University Press.
- Tan, Jacinta, Tony Hope, and Anne Stewart. 2003. Competence to refuse treatment in anorexia nervosa. *International Journal of Law and Psychiatry* 26 (6): 697-707.
- The Mental Capacity Act. 2005. *Office of Public Sector Information*.
http://www.opsi.gov.uk/acts/acts2005/ukpga_20050009_en_1
- The Mental Health Act. 2007. *Office of Public Sector Information*.

http://www.opsi.gov.uk/acts/acts2007/ukpga_20070012_en_1

The Mental Health Act, Department of Health. *The National Archives*.

[http://www.dh.gov.uk/en/Healthcare/Mentalhealth/Policy/InformationontheMentalHealthActs/
index.htm](http://www.dh.gov.uk/en/Healthcare/Mentalhealth/Policy/InformationontheMentalHealthActs/index.htm)