

[Forthcoming in *Inquiry* 52.1 (2009).]

**Mental Capacity And Decisional Autonomy: An
Interdisciplinary Challenge.**

GARETH S. OWEN¹, FABIAN FREYENHAGEN², GENEVRA RICHARDSON³,
MATTHEW HOTOPF¹

¹*Department of Psychological Medicine, Institute of Psychiatry, King's College London, UK*

²*Department of Philosophy, University of Essex, UK*

³*School of Law, King's College London, UK*

Correspondence Address: Gareth S Owen, Department of Psychological Medicine,
Institute of Psychiatry, Western Education Centre, Cutcombe Road, London SE5 9RJ,
UK. E-mail: g.owen@iop.kcl.ac.uk

Abstract

With the waves of reform occurring in mental health legislation in England and other jurisdictions, mental capacity is set to become a key medico-legal concept. The concept is core to the law of informed consent and is closely aligned to the philosophical concept of autonomy. It is also closely related to mental disorder. This paper explores the interdisciplinary terrain where mental capacity is located. Our aim is to identify core dilemmas and to suggest pathways for future interdisciplinary research.

The terrain can be separated into three types of discussion: philosophical, legal and psychiatric. Each discussion approaches mental capacity and judgmental autonomy from a different perspective yet each discussion struggles over two key dilemmas: whether mental capacity and autonomy is/should be a moral or a psychological notion and whether rationality is the key constitutive factor.

We suggest that further theoretical work will have to be interdisciplinary and that this work offers an opportunity for the law to enrich its interpretation of mental capacity; for psychiatry to clarify the normative elements latent in its work and for philosophy to advance understanding of autonomy through the study of decisional dysfunction. The new pressures on medical and legal practice to be more explicit about mental capacity make this work a priority.

Correspondence Address: Gareth S Owen, Department of Psychological Medicine, Institute of Psychiatry, Western Education Centre, Cutcombe Road, London SE5 9RJ, UK. E-mail: g.owen@iop.kcl.ac.uk

Introduction

Mental Capacity is an emerging concept in law and psychiatry. Recent UK law reform has produced major capacity statutes both in England and Wales and in Scotland, the *Mental Capacity Act (2005)* and the *Adults with Incapacity (Scotland) Act 2000*. These formalise the common law on capacity and introduce new standards of explicitness to safeguard patients and professionals aimed at respecting patient autonomy. The older, asylum based, mental health legislation has also undergone reform with mental capacity emerging as a central controversy. In Scotland's *Mental Health (Care and Treatment) Act (2003)* the criterion of 'impaired decision making' was adopted (a notion hard to separate from decisional incapacity). In England the government rejected its expert committee's recommendations to base legal reform on mental capacity and now faces complex relationships between two legal frameworks which can easily apply to the same individual: the amended *Mental Health Act 1983* (not mental capacity based) and the *Mental Capacity Act (2005)* (Owen *et al.* 2008b; Richardson 2007). Controversy continues over whether mental health law should be unified with mental capacity as the explicit basis (Dawson and Szmukler 2006). In the USA, the American Psychiatric Association included mental capacity as a test for civil commitment in 1983 (Stromberg and Stone 1983) and some jurisdictions have adopted this test. In the last decade in the USA there has been renewed concern about ensuring psychiatric research conforms to the law of informed consent and, as part of this, there has been attention to mental capacity assessment in research settings (Carpenter *et al.* 2000). 'Mental Capacity' is a concept that is having increasing practical ramifications.

Yet, for law and psychiatry, mental capacity is a concept which experiences considerable complexity and difficulty when it comes to interpretation: greater clarity is needed. This is especially so because mental capacity is intimately linked to the prized philosophical concept of autonomy. We think that an important first step in taking this call for clarity forward is to journey into the interdisciplinary terrain mental capacity is located and try to identify core issues.

The word capacity connotes an ability, faculty or power to do or achieve something specific. In medico-legal contexts capacity has been defined as the ability to perform a task or make a decision (Culver and Gert 1990). What does 'ability' mean in this context? The ability to 'understand' and 'appreciate' the nature and consequences of a decision is perhaps the standard meaning (Culver and Gert 2004). The ability to make a rational decision is also often added but is more controversial (Culver and Gert 2004). What seems clear is that in medico-legal contexts capacity always relates to some decision, or act. It is only in relation to such a decision or act that the concept gains its moral and psychiatric-legal significance. The bioethicist Carl Elliot (1991), highlighting the moral significance of the capacity concept has put forward the idea that capable decisions have the defining characteristic of being decisions which we may legitimately hold to account. We would, for example, not hold a decision to refuse treatment to account if this decision lacked capacity.

Capacity for decisions, typically, becomes an issue in clinical settings and relates to patients. Who are these patients? There are perhaps two main groups who may require capacity assessment: 'organic' psychiatric disorder (learning disability, dementia and other organic brain syndromes) and 'non-organic' psychiatric disorders (psychotic illness,

depression, anorexia nervosa, etc.).¹ In the former group one might more easily talk of incapacity in cognitive terms (e.g. intelligence or neuropsychological performance too poor to understand a given decision) and in the latter group one may more easily talk of incapacity in non-cognitive terms (e.g., too paranoid to appreciate the significance of an intention to assault an alleged persecutor, too depressed to make a rational decision about suicide). The distinction may help orient the reader and relates to how traditionally in law surrogate decision making in medical settings has been legitimised on the basis of incapacity, understood largely in cognitive terms, whereas surrogate decision making in psychiatric settings has been made on the basis of the risk of harm to the patient or to others, understood largely in value terms, with no reference made to capacity.

The conceptual literature on mental capacity is complex because it mixes philosophical, legal and psychiatric vocabularies. Indeed these three disciplines converge upon mental capacity. For the purpose of explication and clarification, we have found it preferable to try to separate them into three streams of discussion - pointing out areas of contact as we go along. Our aim is to locate the common dilemmas and make some suggestions for future research on a topic which is becoming of pressing practical concern.

I. Philosophical perspectives on capacity and autonomy

Within philosophy and legal theory, the concept of mental capacity for decision-making is intimately connected to that of autonomy. In particular, there is a widespread consensus that one needs to have this capacity at least to a minimal level in order to count as an

autonomous person. For how could one govern oneself without the ability to make decisions? Thus, (young) children, mentally debilitated adults (where debilitation here includes insanity, severe disability, and dementia) and the comatose are not seen as autonomous because they lack the capacity for (rational) decision-making. On the other hand, virtually everyone else is thought of as having the capacity to the minimum level required (Feinberg 1988, p. 28).

However, the consensus ends when it comes to saying what autonomy consists in. Although philosophical debates about autonomy span a number of sub-fields, there are at least two deep fissures between rival and apparently incompatible ways of understanding the nature of autonomy which run through most, or even all, of these debates. The first fundamental disagreement is about whether or not autonomy is value-laden. Thus, we can distinguish between conceptions of autonomy as *moral autonomy* and conceptions of it as *personal, or psychological, autonomy*. On the former view, to be autonomous is to give oneself, or freely subject oneself to, moral norms, values, principles or laws (depending on the view in question). Here the issue is what it is to be self-governing as a moral agent. Personal autonomy on the other hand, is self-government understood in a broader sense: it pertains to persons who govern themselves according to their own reasons or desires, where, depending on the particular view taken, this might include various emotive states, preferences, values, commitments, or character traits. Often accounts of personal autonomy are value-neutral. In this case, a person's own reasons or desires need not be moral ones—he or she can be autonomous *and* non-moral or even immoral. However, there are also accounts which build values or the capacity to appreciate them directly into the conception of personal autonomy, though then the values in questions are not

restricted to moral ones, but can include aesthetic and prudential values or those based on etiquette.

The second important dividing line within philosophical debates about autonomy, which partly cuts across the first one, is the issue of whether and to what extent autonomy must be *rational*. Thus, conceptions of autonomy differ as to whether or not the agent needs to meet certain requirements of rationality, and in their understanding of these requirements (formal/procedural *versus* substantive).

To bring out these disagreements and the difficulties with the respective positions taken on them, we will briefly consider four models of autonomy: (1) the Kantian conception according to which autonomy is moral and rational; (2) the one inspired by H. Frankfurt's work which presents autonomy as inherently neither moral nor rational; and (3) two conceptions of personal autonomy which build rationality into their accounts, one value-neutral and procedural and one value-laden and substantive.

I.1 *Kantian autonomy*

Kant's influential view is an example of autonomy as both inherently moral and rational. According to Kant, autonomy is a property of the will of an agent ([1785], IV:440), who self-legislates moral principles to him- or herself and acts out of respect for them, which, it turns out, is nothing other than being moved by pure practical reason alone. Kant's view is in part motivated by a persistent philosophical problem—how we can be autonomous and responsible if all events in the world (including our actions) are causally determined—and in response to this problem Kant places agents outside of the empirical realm and endows them with the freedom to begin a causal series spontaneously.

There are at least three criticisms of Kantian autonomy, each of which focus on these requirements of rationality and morality. Firstly, the Kantian conception of autonomy is often accused of being too rationalistic, since it portrays emotions as external to the agent and thereby as heteronomous forces which should play no role in moral reasoning and justification. Yet, according to the critics, there are reasons to think that emotions are neither alien to our deepest and truest self, nor unsuitable for practical and moral reasoning or even justification (e.g. Williams 1973, Ch. 13). As we will see below (section III.3), there is also evidence from psychiatry which speaks against a purely cognitive or reason-based model of mental capacity. Secondly, even leaving aside direct criticisms of the strong link with pure practical reason in the Kantian conception of autonomy, this link has been indirectly criticised for coming at a too high cost, since it (allegedly) commits us to metaphysically extravagant views (e.g. again Williams 1973, here p. 228), such as the thesis that agents act outside of time and affect the world on the basis of a form of causation different from the natural one. Finally, critics have objected to the very idea of moral autonomy and especially so to the Kantian version of it (Dworkin 1988, Ch. 3). Among other things, they have argued that even immoral persons are self-governing themselves and should be held morally responsible, but that this would be overlooked if we restricted autonomy to moral autonomy (Feinberg 1988, pp. 43f). Moreover, worries have been raised against such conceptions, since they might (implausibly and objectionably) imply that the state or other individuals can 'force us to be free' or to live according to our 'true self', understood as the morally upright and completely rational Kantian self (Berlin 1958).

This is not the place to discuss the manifold Kantian replies to these criticisms. It suffices to note these three criticisms have fuelled a search for an alternate account of autonomy which would avoid these (alleged) pitfalls. Before turning to the discussion of such accounts, it is, however, necessary to note the following. One important strategy to reply to the second criticism is to emphasise that Kant's theory of autonomy serves a particular kind of purpose, namely, making sense of the practical perspective of deliberation and interaction, and that we can accept a Kantian analysis of this perspective without necessarily having to buy into his conception of transcendental freedom and the intelligible self (e.g. Hill 1992, Ch. 5). From the practical perspective, causal determinism is just not relevant—the fact that I am causally determined to do X does not help me to decide what to do, even if I knew about my being so determined (Korsgaard 1996, Ch. 6). Similarly, by holding others responsible for how they behave we ascribe freedom to them (Korsgaard, Ch. 7; see also Strawson 1962). Thus, it is a presupposition of the practical perspective of first-personal deliberation and second-personal interaction that we think of each other as free, and this presupposition is largely unaffected by whether or not we are in fact determined to act as we do.

However, it should be noted that the practical perspective can break down in particular instances—not for the general reason that the natural world is a deterministic causal series, but rather because of some affliction which impairs the mental capacity of the agent. While the self-ascription and mutual ascription of freedom is the default position, we suspend it with certain human beings, at least on occasion. Thus, in the case of someone who lacks mental capacity, we might upon finding out about this suspend our reactive attitudes to aspects of their behaviour. It might even be the case that from the

perspective of such a person there can be instances in which deliberation can become impossible and self-ascription of freedom is undermined. Contemporary Kantians acknowledge the possibility of the breakdown of the practical perspective (Hill 1992, p. 95). And more generally, they are prepared to admit that our autonomy relies on empirical conditions for its development in and exercise by vulnerable, embodied, socialised and finite creatures such as us (see especially Herman 1993).

Still, there remains a problem here for Kantian accounts of autonomy. Whether these accounts contain any reference to transcendental freedom and the intelligible self or merely rely on explicating the presuppositions of the practical perspective, they are not well-suited to spell out these empirical conditions of agency. For, strictly speaking, there can be no empirical conditions of agency within Kant's theory, since as empirical beings we are mere natural events, pre-determined and in principle predictable, not agents of any sort. Nor could introspection bridge the gulf between the theoretical and the practical perspective—according to Kant, our deepest motivations are opaque even to ourselves (see, e.g., [1785], IV:407). At best, Kantians can suggest a division of labour, where the specification of the empirical conditions is left as a task for the social and human sciences or perhaps to common sense, though even then the danger remains that their inclusion within practical philosophy is *ad hoc*; at worst, the acknowledgement of these conditions is excluded altogether. Similarly, the stark separation of the two perspectives makes it also hard to see how insights from Kant's practical philosophy can play a role in empirical inquiries, such as psychiatry. This is particularly important in the present context, since it also means that the Kantian conception of autonomy seems unsuitable for guiding the clinical or legal practice of dealing with mental capacity. Kantian

autonomy can neither help us with devising an operationalised test for this capacity, nor inform the practical judgements of clinicians. If guidance were offered, it would again run the danger of being *ad hoc* and would have to be based on nothing more than ‘sound common sense’ [*gesunder Menschenverstand*], or on what Kant called in his *Anthropology from a Pragmatic Point of View* [1798] ‘knowledge of the world’ [*Weltkenntnis*].”

I.2 The Frankfurt-inspired model: autonomy as reflective endorsement

One prominent view among the theorists who argue for personal autonomy are higher-order accounts inspired by H. Frankfurt's influential discussion of freedom of the will (Frankfurt 1971). On these accounts, we distinguish between first-order desires (pro-attitudes, preferences, commitments, values) and second-order desires (attitudes, etc.), that is, between wanting something and wanting to want it.ⁱⁱ Autonomy is then understood to require second-order identification with one's first-order desires. Conversely, one is lacking autonomy if one wants to do something which one does not want to want to do. Here the standard example is of an unwilling addict, who desires to take a drug, but does not identify with this desire and would prefer to be without it. Moreover, there is an interesting further case, namely, of someone who has no second-order desires at all. In this case, the individuals (whom Frankfurt calls 'wantons') lack autonomy by default—given the structure of their desires (or rather, the lack of structure), they could not be autonomous. One common way of framing the key idea of this model is in terms of 'authenticity' (Dworkin 1981): a desire, value, commitment, or character trait is authentic (or genuinely one's own), if one reflectively identifies with it.

This account of autonomy conceives of it neither in moral nor rational terms. While identification might acquire moral relevance in a number of ways—for example, it might suffice for moral responsibility—there is no internal link to morality implied in this structural constraint on an agent's motivational set. Thus, even someone who identifies with and endorses his or her immoral first-order desire counts as autonomous on this model.ⁱⁱⁱ Similarly, autonomy does not necessarily involve meeting any rationality requirements—the endorsement of first-order desires at a second-order level can be a wholly non-rational process. In fact, it need not even involve deliberation (see already Frankfurt 1988, p. 22). Proponents of this model want to avoid over-intellectualising autonomy and admit that second-order identification might be achieved even when the person in question did not consciously attend to the matter, but would endorse his or her first-order desires if he or she had attended to them (see also Dworkin 1988: 17).

One advantage of the Frankfurt-inspired model is that it seems to avoid the metaphysical commitments of many traditional models, including Kant's. It analyses autonomy (and personhood) in terms of something which is (metaphysically speaking) innocuous, namely, in terms of a psychological capacity of structuring one's desires in a particular way. By introducing the idea of a hierarchy of desires we can explain the resolution of conflicts among first-order desires by appeal to a 'true self', but without commitment to a homunculus or transcendental self (Bratman in Taylor (ed.) 2005, p. 37). Other advantages are that this model need not conceive of emotive states as heteronomous, alien influences and that it frees autonomy from overly evaluative and moralised presuppositions. For these reasons, the Frankfurt-inspired model may seem

more promising for deployment in clinical and legal use than the Kantian account of autonomy.

However, there is also a problem with this model. The Frankfurt-inspired model of autonomy has been criticised in a number of ways, ranging from doubts expressed about the special authority of second- and higher-order desires (Watson 1975, pp. 217-219) to worries that this model leads to an infinite regress of levels of identification. While ultimately all of these objections suggest that identification with first-order desires is insufficient for autonomy, this point is perhaps best brought out in cases of manipulation, brainwashing or indoctrination. These cases—such as someone being pressurised into a decision—are often thought to be paradigmatic examples of the absence (or violation) of autonomy. It would thus be a fundamental failing of the Frankfurt-inspired model, if it lacked the resources to account for them.

Frankfurt and those defending similar views have variously tried to shore up the authority of the higher-order endorsement to answer this criticism, either in terms of the decisiveness of the identification (Frankfurt 1988, p. 21 and Ch. 12; for further refinements see Frankfurt 1999), or by thinking of the reflective level in terms of self-governing policies, rather than just as another layer of wants (Bratman 2007). However, none of these refinements really seem to fully exclude the possibility of manipulation (Taylor (ed.) 2005, pp. 4-10).

The philosophical discussion of manipulation cases often involves toy-examples. Even so, the issue they raise might make one sceptical about putting the Frankfurt-inspired model to use in the clinical or legal contexts of deciding about mental capacity. The danger of being manipulated or pressurised into a decision (for example by family

members or better informed and paternalistic practitioners) lurks here too—in fact, there already has been one important legal case where outside or ‘undue’ influence on the patient was judged relevant to decision-making capacity (*Re T, 1993*).

I.3 Personal autonomy and rationality

In order to account for the manipulation cases, one might depart more radically from the Frankfurt-inspired model of personal autonomy, for example by adding a second condition on top of the identification/authenticity condition. We will consider two versions of such a departure, starting with the one which adds (or substitutes) a procedural condition.

The most prominent exponent of this model is Gerald Dworkin. Like Frankfurt, Dworkin operates with the idea of first- and second-order desires and, at least pre-1988, makes reflective endorsement a condition on autonomy. However, Dworkin adds a second condition which constrains the way identification with first-order desires can take place in order to still count as autonomous (Dworkin 1981): if the process was affected by autonomy-undermining influences, then the identification is not counted as conferring autonomy. By adding this requirement of ‘procedural independence’ to the account, the model of personal autonomy seems to answer at least some of the objections made to the Frankfurt-inspired model. Thinking back to the example of a person pressurised into a decision, Dworkin could agree that this is not a case of autonomy, since undue influence violates procedural independence.

In fact, in his later writings, Dworkin moves even further away from Frankfurt by dropping identification altogether. Instead, his view becomes that autonomy requires that

we can critically reflect and change our desires (albeit not at will, but only for good reasons) and that this ability for critical reflection meets the procedural independence criterion (Dworkin 1988, pp. 15-17). Part of the reason for this change is Dworkin's appreciations of the manipulation or indoctrination cases: for it is not identification which is interfered with in such cases, but the capacity to critically reflect on the desires and values in question, and this suggests that autonomy, which we intuitively think is interfered with here, centrally consists in something else than identification.

The crucial challenge for this model of personal autonomy is to identify which influences are autonomy-undermining and which are not, and to identify them on the basis of a procedural account alone. A lot of the further refinement of this model has been concentrated on working out what procedural independence requires, for example, in respect with preference-formation (see Christman 1991). If the defenders of this model are to make good on its promise, then the standard of procedural independence needs to capture all of the intuitively non-autonomous influences without references to objective reasons or the rationality of outcomes. Only then can they lay claim to content-neutrality, which would have the advantage of making autonomy compatible with the entrenched value pluralism of modern liberal societies (see, e.g., Christman in Taylor (ed.) (2005), Ch. 12).

However, it is far from obvious that procedural independence suffices to identify autonomy-undermining influences and conditions. The standard of rationality employed here is minimal and only concerns the process of reasoning—beliefs and desires need to be in some sort of coherence and this coherence should not be the result of factors which inhibit or outweigh reasoning processes in a problematic way (with the paradigm

examples being manipulation by others and lack of certain cognitive faculties). Yet, there are some cases in which such a minimal, procedural standard of rationality is met, but we would still not want to call the person (or their specific decisions) autonomous. Thus, to anticipate the psychiatry section, there may be some patients who, while otherwise in possession of their cognitive abilities, have a set of delusional or false beliefs and/or distorted values who reason as well on the basis of these beliefs or values and might have as coherent a set of beliefs and desires as non-pathological persons do and have. On the procedural rationality model these patients would count as autonomous (and might, hence, be granted the full legal right of self-determination). Yet, this seems counterintuitive (and also contravenes psychiatric practice).

Considerations such as these might motivate the turn to yet a third model of personal autonomy, according to which we need to add an externalist criterion: responsiveness to 'objective reasons'. Such reasons might be of an epistemic nature, so that responsiveness to them implies that one is not completely ignorant or has no manifestly false and delusional beliefs (e.g. Berofsky 1995). And these reasons can also involve orientation by and appreciation of values (e.g. Benson 1987). If so, then personal autonomy gains an inherently normative, even moral, dimension. Thus, Meyers, among others, claims that the conduct of the agent needs to be within the bounds of the moral permissible to count as (genuinely) autonomous (1989, Ch. 1).

This threatens to bring us back full circle: by building rational and normative standards into conceptions of autonomy, we seem to settle ourselves with some of the same problems as the ones faced by the Kantian view. In particular, difficulties arise because there is deep, and arguably, reasonable pluralism about values and the nature of

substantive rationality within modern societies. Conception of autonomy which contain a substantive rationality standard will thus face the challenge of being sufficiently broad to be compatible with this pluralism, but also sufficiently fine-grained to be of use in identifying lack of capacity in legal and psychiatric contexts. Moreover, such conceptions are in tension with an intuition which is widely accepted (at least within the context of modern, liberal societies), namely, the intuition that most adult human beings are self-governing their lives to an extent that they should be granted the full right of self-determination, despite the fact that they often do not meet standards of substantive rationality (and perhaps even when they fall short of the minimal or procedural standards of rationality). In fact, as the law section of the paper notes, the protection of the autonomy of individuals within liberal states includes the protection of irrational decisions.

Secondly, integrating a normative dimension into accounts of personal autonomy risks collapsing personal autonomy into moral autonomy, and thereby making it subject to the same worries as those which are raised against the moral aspect of Kantian autonomy, especially the danger of legitimising paternalism and forcing people to be free (see Christman 1989, 12).

One might reply that a move away from a morally neutral conception of autonomy is, actually, a good thing, despite the worries considered earlier. For example, as especially Kantians note (e.g. O'Neill 2003), if personal autonomy is understood in a morally neutral way, then it is unclear whether it could play the important moral (and legal) roles which autonomy is meant to play, such as justifying the right of individuals to determine their own lives. Something which is not itself inherently moral might not be

the right kind of value to play such roles. Alternatively, it has been suggested that personal autonomy, while different from moral autonomy, only has value if it is employed in morally permissible ways and in the pursuit of the (moral) good (Raz 1986, Ch. 14).

Use of conceptions of autonomy which build in rationality of either a procedural or substantive kind seems promising for establishing mental capacity, since these conceptions tend to do without strong metaphysical commitments, but offer more guidance than Frankfurt's model. Still, for the reasons pointed out, these conceptions also face problems. In particular, we have seen that they do not fit actual practices in the clinical or legal contexts, since delusional patients are not treated as autonomous even when procedurally rational and (substantively) irrational decisions can be protected by the law. Thus, before either these conceptions can be utilised in the clinical or legal context, further refinements both within and between the three areas will be required.

In sum, debates about autonomy in philosophy can usefully be characterised with reference to the two related issues of whether or not (a) autonomy is morally (and evaluatively) neutral in nature and (b) it is inherently constrained by rationality of some sort. There are good reasons to suspect that the same tensions (or, at any rate, very similar and related tensions) also appear in the legal and clinical contexts, and we now turn to the discussion of mental capacity and autonomy within these other two contexts.

II. Legal perspectives on capacity and autonomy

Contemporary human rights doctrine accords great significance to the principles of self-determination and respect for autonomy and medical law tends to reflect these principles. The question of why liberal democracies place so much importance on self determination goes far beyond the scope of this paper, but in the medical context consent is the legal tool used and the weight this places on the concept of mental capacity is the subject of the following discussion. In relation to medical treatment the law requires the patient's consent and will respect a decision to refuse medical treatment, to do otherwise would be to deny the patient's right to self-determination. The requirement of consent thus provides the mechanism through which the patient protects her autonomy and right to self determination. But to attract the law's respect the refusal must be made with capacity because only such a decision can be regarded as autonomous. At this point the law vests no 'moral' significance in the notion of capacity itself. It is seen simply as a necessary condition for the valid exercise of self-determination. There are, inevitably, cases where unswerving respect for individual autonomy gives rise to 'moral' difficulties, where a mother's refusal of a c-section endangers the unborn child (Re MB) for example, or where a mature minor refuses life-saving treatment (Re W and Re P), but even here the courts adhere to the rhetoric of individual autonomy. It is perhaps in the assessment of capacity itself that moral values exercise a greater influence (see below and R v Collins ex p Brady).

The law also operates a presumption in favour of capacity. An adult will be presumed to have the capacity to consent to or to refuse treatment unless there is evidence to the contrary. The concept of capacity is therefore crucial to the law of consent and the

courts have to devise criteria according to which they can establish the presence or absence of capacity in cases where the initial presumption is challenged. Obviously the courts are not alone in having to devise criteria for the recognition of capacity. As will be seen below clinicians have also to struggle with the concept and tools have been devised to assist them in the assessment of capacity (Hotopf 2005). For both lawyers and clinicians the concept of capacity is continuing, it is seldom entirely present or entirely absent, but for both disciplines it becomes necessary to determine whether it is present or absent in relation to specific decisions. In both law and clinical practice capacity becomes a binary concept only at the point of decision-making. Lord Donaldson's remarks in *Re T* (1993) illustrate the process:

“Doctors faced with a refusal of consent have to give very careful and detailed consideration to what was the patient's capacity to decide at the time when the decision was made. It may not be a case of capacity or no capacity. It may be a case of reduced capacity. What matters is whether at that time the patient's capacity was reduced below the level needed in the case of a refusal of that importance (116)”

II.1 The legal criteria

Unfortunately there are no universally agreed criteria for the recognition of capacity operating across all legal jurisdictions but most systems do adopt a functional, as distinct from a status, test in relation to medical treatment. Capacity is assessed in relation to each decision, its presence or absence is not simply assumed on the basis of the individual's status.

More generally, Roth (1977) and Berg et al (1996) have analysed the tests for capacity which have emerged from the (mainly US) case law. In relation to treatment decisions these tests or criteria have been refined by Grisso and Applebaum (1998) as:

- 1) The ability to express a choice about treatment,
- 2) The ability to understand information relevant to the treatment decision,
- 3) The ability to appreciate the significance of that treatment information for one's own situation,
- 4) The ability to reason with relevant information so as to engage in a logical process of weighing treatment options.

While Grisso and Appelbaum's analysis is designed as a descriptive definition written to inform clinical practice, rather than as an authoritative statement of law, it provides a useful background against which to consider the position as it is evolving in UK, primarily English, law.

In the UK the concept of capacity in relation to medical treatment was initially developed by the common law, or case law, and there was no relevant statutory framework. Then in 2000 the Scottish Parliament passed the Adults with Incapacity (Scotland) Act and in 2005 England and Wales followed with the Mental Capacity Act (MCA). According to the MCA a person lacks capacity in relation to a particular matter if he is unable to make a decision for himself because of "an impairment of, or a disturbance in the functioning of, the mind or brain" (section 2(1)). A person "is unable to make a decision for himself if he is unable:

- a) to understand the information relevant to the decision,
- b) to retain that information,

- c) to use or weigh that information as part of the process of making the decision,
- or
- d) to communicate his decision” (section 3(1)).

To a large extent this statutory definition builds on the position at common law (Law Commission 1995) and it is likely that the courts will refer to the existing case law when it comes to applying the statutory words. Any discussion of the definition of capacity in English law therefore requires reference to both the common law and the MCA.

A comparison of Grisso and Appelbaum and the MCA immediately indicates that both definitions include the ability to express a choice or to communicate a decision. Such a requirement is seldom expressed in the common law but can be implied. Both definitions also include the ability to understand information relevant to the treatment decision, while the MCA and the common law specify in addition the ability to retain that information (Re C 1994 and Re MB 1997). While these two requirements, communication and understanding, might be most easily interpreted in cognitive terms, i.e. attention, memory, and language, notions of relevance and even communication itself bring their own areas of ambiguity (Manson & O’Neill 2007).

In relation to Grisso and Appelbaum’s third and fourth criterion, the abilities to appreciate and to reason, more serious dilemmas emerge, particularly in the context of mental disorder and the relationship between capacity and rationality. Appreciation and the ability to reason are avoided in both the MCA and the English common law where to “use” and to “weigh” are preferred. But, whatever the precise terms used, the underlying difficulties remain.

Appreciation can be seen as an attempt to capture the requirement that the person should be able to grasp the relevance of the information to their own situation. In this sense it is similar to the clinical concept of insight, discussed below. It is evident that a notion similar to appreciation is reflected in the English law requirement of “use”, or even “understand” (*Trust A and Trust B v H* (2006), *R (B) v Dr SS* (2006))

Grisso and Appelbaum’s fourth criterion, the ability to reason with the relevant information so as to engage in a logical process of weighing treatment options, raises perhaps the most fundamental question: to what extent is rationality or the ability to reason logically related to the legal notion of capacity? At the level of rhetoric at least, English law is prepared to recognize capacity despite the presence of irrationality in either the decision-making process or the outcome. In relation to the process:

"A mentally competent patient has an absolute right to refuse to consent to medical treatment for any reason, rational or irrational, or for no reason at all, even where that decision may lead to his or her own death." (*Re MB*, 1997)

And again:

“the patient’s right of choice exists whether the reasons for making that choice are rational, irrational, unknown or even non-existent.” (*Re T*, 1993).

So, the absence of rationality in the decision-making process cannot alone amount to incapacity (*Gunn et al.* 1999). Neither can the irrationality of the outcome. In the quote just given Lord Donaldson goes on to state:

“That his [the patient’s] choice is contrary to what is to be expected by the vast majority of adults is only relevant if there are other reasons for doubting his capacity to decide.”

Thus at the level of rhetoric English law reflects a generous and liberal approach to self-governance by refusing to afford priority to any one understanding of rationality.

However, despite denying that irrationality can on its own amount to incapacity, the courts are quite prepared to accept that it can be a symptom or evidence of incapacity (*Re MB*). It is important here to remember the role which the capacity/incapacity distinction performs in law. It has evolved as a means of determining which decisions are to be regarded as autonomous and therefore attracting respect, and which are not. Thus if the requirements of capacity are too generous there is a risk that decisions which are not truly autonomous will be protected, while if they are too strict some truly autonomous decisions may be inappropriately ignored (Gunn et al 1999, Kress 2004). (See the discussion in the philosophy section).

Mirroring the third concept of personal autonomy outlined in the philosophy section - with its introduction of 'objective reasons' - consideration has been given to a legal criterion for capacity that the decision maker be able to provide "recognizable" (Freedman 1981) reasons for the decision.. Such a criterion concentrates on the reasoning process preceding the decision, but goes beyond the requirements of procedural rationality, since it would cover decision-making based on patently false premises or distorted values. To some extent such an approach is already reflected in the English case law. In *Re T* (2004) the court was clear that a refusal of a blood transfusion based on the patient's belief that her blood was evil should not be respected. The patient's misperception of reality had rendered her incapable of using and weighing the relevant information. She had no 'objective reason' for her decision. Then in *Trust A and Trust B v H* (2006) the court explained that:

“a compulsive disorder or phobia may prevent the patient's decision from being a true one, particularly if conditioned by some obsessional belief or feeling which so distorts the judgment as to render the decision invalid.”

While both these cases can be seen as examples of the law's concern with the process of and input into decision-making rather than the outcome - the decision-making lacks capacity because it is based on a misperception of reality, not simply because it is an “irrational” decision - the distinction between process and outcome can become very fine. If we have capacity the law will respect our refusal to accept treatment however irrational that refusal may be. So in English law an extraordinary or “stupid” refusal can still be capable if there is no other evidence of impaired decision-making. Distorted reasoning can, of course, supply that evidence, as in *Re T* (2004). But in cases where the reasoning process is never revealed the law's presumption in favour of capacity might be thought to drive it to an over generous attribution of capacity, requiring it to respect decisions which are not truly autonomous. However, English case law has consistently required the level of capacity to relate to the gravity of the decision in issue (*Re B*, 2002). In *Re T* (1993) Lord Donaldson emphasised the requirement that the relevant capacity must be “commensurate with the gravity of the decision”:

“[w]hat matters is whether at that time the patient's capacity was reduced below the level needed in the case of a refusal of that importance, for refusals can vary in importance. Some may involve a risk to life or of irreparable damage to health. Others may not.”

Thus despite English law's apparent generosity towards the irrational decision, the gravity of the consequences flowing from such a decision will affect the rigour with

which the courts scrutinise the reasoning leading up to it. Matters of value, via the concept of “gravity”, can be seen to frame capacity assessment. Here we see that the law is struggling, like philosophy, over the question of how moral or evaluative the concept of mental capacity may become and over how to interpret a rationality criterion.

III. The psychiatric perspective on capacity and autonomy

Psychiatrists approach patient autonomy from the principle of duty of care and in settings where there is mental disorder, full autonomy will not be simply assumed. This initial starting point makes the psychiatric perspective on individual autonomy different from the legal perspective with its presumption of capacity. Whilst clinicians face the moral obligation to respect autonomy they also face the obligation to enable it where it is compromised by psychopathology. This involves surrogate treatment decision making and historically has been performed in a graded fashion without explicit reference to the legal concept of capacity. Some clinicians have argued that this practice, which may be construed as paternalism, ‘maximizes’ patient autonomy (Wulf 1995; Komrad 1983).

Clearly, psychiatrists do not take on duties of surrogate decision making for all their patients so the question arises: what are the relevant clinical factors? This question has a similar form to the legal question: which factors are relevant in deciding upon legal capacity? As we have seen above the legal answers to this question have not been straightforward and have mirrored dilemmas that exist within the philosophical analysis of autonomy. We will now look at the psychiatric answers. Four areas stand out in the psychiatric literature: insight and neuropsychological functioning; disorder of thinking

(which incorporates delusion); disorders of emotion; risk. We will address them sequentially.

III.1 Insight and Neuropsychological function

Insight is a complex phenomenon which is core to clinical concepts of schizophrenia (McEvoy *et al.* 1989) but also to manic episodes of Bipolar Affective Disorder and forms of organic brain injury (Amador 2004). It is a term that is easier to use than to define (Fulford 2004) but the basic concept – that of an *unawareness of illness* – has acquired an acknowledged place in psychiatric assessments (David 1990). The concept has been much discussed in the setting of patients with “anosognosia” following damage to the non-dominant cerebral hemisphere in which frank denial of paralysis of the left side of the body can be evident. The significance of loss of insight for decision-making capacity can be seen if one considers patients who are in the following predicament:

- 1) They are ill
- 2) They are unaware of this illness (they lack insight).
- 3) There is a causal connection between their illness and their unawareness of this illness.

In this situation the patient’s decision-making is caught in a loop whereby to acknowledge illness he or she/the patient needs to decide to have treatment but to validly consent to treatment he or she needs acknowledgement of illness.

Kress (2004) has argued that advances in our understanding of low insight, including its neural basis, provide good reason for giving it a more significant role in legal evaluations of capacity. It is true that poor insight has been strongly associated with

incapacity in patients with psychiatric disorders (Cairns *et al.* 2005; Owen *et al.* 2008a) but two things are worth clarifying at this point. The first is that insight is a series of judgements made within a clinical interview rather than an impersonal psychological test and depends upon a prior assessment that the patient has an illness. The second is that in severe psychiatric illness low insight has weak association with neuropsychological impairments (Aleman *et al.* 2006). So, in psychiatric settings, one needs to be cautious about attributing to insight the status of a psychological or neuropsychological variable – it has an inherently normative element which results from it being a judgement about the patient’s ability to recognise certain experiential states as pathological. And such judgements are value-laden

Perhaps because of this concerns have been expressed over the clinician-centric character of the concept (Perkins and Moodley 1993) and the psychiatric rather than legal origins of this construct (Diesfeld 2003). Certainly, “having insight” must mean more than “agreeing with the doctor” and insight, given its ethical and legal significance, needs to be transparent to ethics and law as well as to psychiatry.

General neuropsychological measurements of attention, memory function and language such as the mini mental state examination (Folstein *et al.* 1975) have been studied as risk factors for incapacity. In settings where organic mental disorders predominate such variables have shown strong and consistent associations with incapacity (Okai *et al.* 2007). In settings where non-organic mental disorders predominate, however, the associations have been weak and less consistent (Okai *et al.* 2007). More detailed neuropsychological modelling of capacity, including measurement of ‘executive’ or frontal cortex functioning, has been undertaken by Marson and

colleagues in an organic setting (Marson and Harrell 1999a; 1999b) and by Palmer and colleagues in non-organic settings (Palmer *et al.* 2004). Marson *et al.* have used Alzheimer's disease as a target disorder. Here, the inability to meet a 'rational reasons' standard (*a standard accepted in many US jurisdictions*) correlates with difficulties in executive tasks (Marson *et al.* 1995). Such findings, though strictly speaking limited to Alzheimer's disease, have seeded notions linking neural systems underlying executive and frontal function to legal capacity (Dunn *et al.* 2007; Marson 2001). Palmer *et al.* have used schizophrenia as their target disorder and have reported weak associations between a variety of neuropsychological impairments – not limited to executive functioning – and incapacity.

Marson and Palmer's work can be seen as an instance of what has become known as "neuroethics" – a heterogeneous discipline which addresses the interface between neuroscience and ethics (Illes 2006). One view of neuroethics is that it reconstructs ethical/legal categories as neuroscientific ones (Churchland 2006). On this view the distinction between capacitous and non-capacitous decisions is a distinction between certain groupings of neuroscientific variables. A capacitous decision is a decision made by a brain functioning within a certain set of neurochemical or neurocognitive parameters; an incapacitous decision is a decision made by a brain operating outside of those parameters. The difficulty of this approach, aside from questions about empirical strength, is that in practice the neuroscientific parameters will be derivative from prior ethical and legal norms rather than the other way around (Buller 2006). At the level of individual clinical assessment neuropsychological performance, for example, becomes relevant to capacity determination only when it is meaningfully related to the abilities

considered relevant for a decision-making situation (Marson 2001). Neuropsychological test results, on their own, even in organic brain syndromes conceptualised neuropsychologically, do not determine capacity judgements. This applies even more strongly in non-organic syndromes which are not conceptualised neuropsychologically. So neuropsychological modelling of mental capacity, like insight, exists within a normative, or evaluative, context.

III.2 Disorder of thinking

Links between insanity and irrational thinking are recurrent themes in the history of psychiatry. The prototype here is delusion (Gert 1995) which is a key criterion for the diagnosis of schizophrenia and related disorders. Delusions are associated with decision-making incapacity (Cairns *et al.* 2005; Owen *et al.* 2008a). They are understood by psychiatrists within a descriptive language which has been difficult to satisfactorily operationalise (David 1999). Conceptualisations which have been proposed include: A false, fixed belief based on incorrect inference about external reality and out of keeping with that person's culture (American Psychiatric Association 2000); judgements held with an incomparable subjective certainty, imperviousness to other's experience and to compelling counter-argument and with content that is impossible (Jaspers 1963) and imagined anomalous versions of the world (Cutting 1997).

If delusions are impairments of rational thinking are they impairments of formal or informal reasoning? Several studies in experimental psychopathology have sought to investigate the relationship between delusions and formal modes of reasoning using both syllogisms and probability theory as formal standards. These studies have not found

impairments of formal reasoning in patients with delusions compared to healthy controls (Cutting 1997; Kemp *et al.* 1997). On conditional probability tasks people with delusions have a tendency to be less conservative in their inferences than healthy controls in a manner which actually brings them into greater conformity with Bayesian norms (Garety and Hemsley 1997). There is also some evidence that when syllogistic deductive reasoning is pitted against informal reasoning people with schizophrenia actually do better than healthy controls on formal reasoning (Owen *et al.* 2007).

Results such as these pose difficulties for efforts to develop a rationality standard as the basis for deciding decision-making capacity. Concerns about a rationality standard have also been raised in relation to anorexia nervosa where, using in-depth interviews, the argument has been made that anorexic patients maintain formal reasoning but do so in the context of disordered valuations (e.g. to be thin is more valuable than to live) (Tan and Hope 2008). Moore (1984) argues that if behaviour has become grossly inconsistent, i.e. where actions repeatedly contradict stated goals then we lose grounds for attributing autonomy to a person and they cannot be held capable to make decisions. But this puts a lot of weight on a formal reasoning standard – consistency – and schizophrenic patients with delusions and patients with anorexia nervosa do not demonstrate impairment here. Thus a formal rationality standard would be charitable to delusions and anorexia nervosa. This would seem to run counter to the clinical consensus that delusions and anorexia often impair capacity. Perhaps they do this because they impair substantive rationality or values rather than formal reasoning.

III.3 Emotion

All psychiatrists will have experience of patients who when in states of severe depression, anxiety or elation make decisions which, on hindsight, they regard as decisions they regret, did not control (often dramatically so) and wish to be protected from making again. In the context of illnesses such as unipolar or bipolar affective disorder clinicians will commonly override the treatment wishes of patients in such states and treat compulsorily if necessary.

There has been concern however from some clinicians that these emotional influences on decision-making are not sufficiently appreciated because the standard approach to capacity is oriented around cognitive or non-emotional features of the mental state (Bursztajn *et al.* 1991; Elliot 1997; Rudnick 2002; Culver and Gert 1990, 2004). These authors argue that altered emotional states 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.

Cases of severe depression where the legal criteria for capacity have apparently been met, but where the patient’s values or preferences seem radically disturbed, have generated some peculiar phraseologies such as ‘competently made irrational decisions’ (Sherlock 1983). The danger here is that patients can be erroneously judged to have capacity because of an inadequate concept of capacity (Culver and Gert 2004). Moorhead and Turkington have called for the role of emotional capacity to be clarified to avoid these types of error (Moorhead and Turkington 2002). How this is to be achieved remains a challenge.

Another way of seeing the role of emotion in clinical capacity assessments is to consider the situation when emotions are unimpaired in the setting of clear cut impairments of cognitive function. Such situations can be seen in patients suffering from dementia or learning disability. Dubler (1985) has argued that it can be appropriate to consider even severely demented patients as retaining capacity. She maintains that strong preferences often survive cognitive deficit and can refer to a documented pattern of preferences and a clear value system – what she calls “sedimented life preferences”. She writes that “when this sort of theme emerges, buttressed by surprisingly articulate statements (such as those describing the death of loved ones at home) and supported by confirmatory statements of family and friends, it deserves great weight and respect; in some cases, ... it should be permitted to override issues of present incapacity” (Glass 1997, p. 29). This suggests that certain expressed preferences in settings of marked cognitive impairment yet normal emotional functioning should be considered capacitous. This would be consistent with the law’s decision-specific concept of capacity and with the clinical practice of assessing capacity even when a diagnosis of dementia or learning disability is established. It indicates that emotion, rather than procedural or substantive rationality, may be the relevant capacity consideration in some settings. This puts pressure on philosophical concepts of autonomy emphasising a rationality criterion.

III.4 Risk-relativity

Buchanan and Brock (1989) formulated the idea that the standard of capacity testing should be adjusted according to the risks of the decision being made. On their account, a patient who refuses life-saving treatment, for example, should have to pass a harder test

of their capacity than if they accepted it. This is a risk-relative or sliding scale concept of capacity testing which aims to protect patients or others from harm. It views capacity as functionally tied to risk and it allows there to be a distinction between capacity to consent to treatment and capacity to refuse treatment (Wilkes 1999).

Others (Wicclair 1991; Cale 1999; Demarco 2002) have argued that justifying the risk-relativity of capacity assessments on the grounds of protecting patients or others from harm does violence to the concept of autonomy. Instead of tying capacity to risk these authors tie risky decisions to high cognitive demand and argue that it is for this reason that a more stringent level of capacity be expected for patients who want to do something risky – risky decisions are just cognitively harder decisions. The authors argue that this formulation respects patient autonomy and capacity is kept conceptually distinct from risk.

The difficulty with this account is whether decisions concerning death or decisions which carry dangers for others (decisions usually regarded as risky) are *necessarily* cognitively harder than decisions which are of low risk (Buller, 2001). We can readily acknowledge that they are more profound but they may require less manipulation of information. It may, for example, require less manipulation of information to know whether to carry a knife or refuse chemotherapy in advanced cancer (risky decisions) than to know whether to take the train or the bus to work (low risk decision).

Another way of formulating risk relativity is to highlight that capacity determination, like any test, carries with it an error margin. On this view, if a patient decides on something risky, then this error margin starts to become significant in a new

way (Demarco 2002). This is because in a risk situation 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. In practice this involves seeking more information of relevance to a capacity judgment and this may take time (Buchanan 2004). To obtain this time, it may only be practicable to detain the patient compulsorily. Buchanan (2004) points out that this approach to capacity – 'a leaving room for error' approach – was favoured by the President's Commission for the Study of Ethical Problems in Medicine (President's Commission 1982).

Either way the psychiatric assessment of mental capacity has not been separated from risk and the values which go with risk. Like the law's insistence that the standard of the capacity assessment should be commensurate with the gravity of the decision, the psychiatric assessment of capacity is carried out within a value context.

This is significant in two respects: firstly, it means that if psychiatry aims at a completely value-neutral or even value-free conception of mental capacity it will come unstuck. Secondly, it means that if philosophy or law seeks value neutrality for mental capacity by deferring to the psychiatric perspective there will be disappointment. As seen, the normative and evaluative nature of psychiatric assessment affects all four of the central areas discussed here. In fact, when it comes to psychopathological phenomena such as delusions this is so in more than one way: psychopathologists rely on normative

considerations not just in judging the significance of a delusion for mental capacity but in order to *identify* delusions. This brings the psychiatric perspective on mental capacity even closer to considerations of substantive rationality and valuation and calls for a much more systemic examination of the relationships.

Conclusion

We have crossed 3 perspectives which converge upon autonomy and mental capacity: philosophical, legal and psychiatric. Each perspective approaches autonomy and capacity in its own terms and, at times, the observer of this terrain might feel that 3 separate discussions are being had. In this paper we have highlighted points of contact between the perspectives along the way but, more importantly, have argued that each perspective, within its own terms, struggles with the same dilemmas: whether autonomy and capacity should be a moral (or evaluative) notion and whether some kind of rationality is the key constitutive factor.

We have suggested that influential philosophical concepts of capacity and autonomy, when moral (e.g. Kantian), risk detachment from legal and psychiatric decision-making and when non-moral (e.g. Frankfurtian) risk failure to supply the normative elements which legal and psychiatric perspectives require. With the rationality dilemma, we have suggested that when concepts of autonomy and capacity adopt a strong rationality standard they risk unwelcome confrontation with legal and psychiatric reality and when weak they risk leaving law and psychiatry without adequate normative orientation.

Much extant thinking about these matters may be guided by the all too familiar ought/is and fact/value distinctions. In the area of autonomy and mental capacity we suggest that these distinctions are at very high risk of fundamentally distorting the practical application of mental capacity in concrete medical contexts. Kantian notions of autonomy and capacity that discount legal cases and psychopathological ‘facts’ risk leaving legal and medical decision-making about mental capacity almost entirely unaccounted for. Personal, or psychological, notions of autonomy and capacity that discount the substantial normative weight that falls upon the legal concept of capacity and veil the normative elements within psychopathology itself risk bringing about a situation in which both anti-psychiatry (e.g. Szasz 1999) and pro-psychiatry movements (e.g. Fuller-Torrey 2008) prosper with all the politicizations that ensue. Szasz thinks psychiatry and mental health laws should be abolished because mental disorders are normative; Fuller-Torrey thinks that psychiatry and mental health laws should be strengthened because mental disorders are brain diseases and morally neutral.

In this arena, a retreat to either philosophical positions on autonomy and capacity completely detached from actual practices and facts on the ground or to medical positivist positions is likely to prove counter-productive. We seem to be left with inescapable mixing of ought and is and inescapable interdisciplinarity. We suggest that further theoretical work will have to bring philosophical thinking about autonomy and capacity into direct contact with psychopathological cases in which mental capacity is in doubt. There is a need to investigate in more detail what the normative elements of mental capacity assessments are and what implications in-depth study of psychopathological cases may have for the law and philosophy.

In depth case studies of the impact delusions, severe mood disturbances and frontal/dysexecutive brain syndromes have on decision-making capacity is an example of how this research might proceed. In each case, there would be analysis of the distortions of the decision making process as the patient encounters it and thereby analysis of the distortions of decision-making rationality or evaluation in these disorders. This analysis is needed because both philosophy and law have come to place considerable weight on the *process* of decision-making. Philosophy has the challenge to identify which influences are autonomy-undermining and which are not but too often the testing of philosophical models of autonomy have relied on tinker-toy examples constructed from the armchair which are unrecognised by practicing psychiatrists. Law has the intricate task of assessing the capacity of a decision-making process even where the resulting outcome is sharply at odds with medical recommendations but the case law on capacity is fraught with internal tensions which may leave the practitioner feeling confused. If philosophy and law are to make good on their emphasis on the decision making process rather than the outcome both will require a richer understanding of decision function and dysfunction. Such understanding will assist the law's interpretation of existing legal criteria or provide stimulus for their reform; it will also assist in advancing philosophical concepts of autonomy. There remains the dream of a purely neuroscientific account of decision function but this is both too far off on the horizon to be useful to us here and now and also not suitable to capture the normative commitments that shape the legal and clinical domains. What is needed now are new efforts in case analysis in which philosophical and legal models of autonomy and capacity are tested against real cases of incapacity encountered in the clinic, and in this way also helping to clarify just what forms of

neutrality can and should be preserved in the assessment of capacity. This case analysis will require a type of patient interview that explicitly tries to probe the decision-making process as the patient experiences it. These new efforts might be called ‘philosophical psychiatry’ and relate to old lines of inquiry in the tradition of philosophical anthropology. The new pressures on medical and legal practice to be more explicit about mental capacity make this work a priority.

Acknowledgements

We would like to thank Peter Hayward, Jules Holroyd and the editor of this journal, Wayne Martin, for comments on earlier versions of this paper.

ⁱ The terms ‘organic’ and ‘non-organic’ are psychiatric terms of art rather than connoting anything specific about the mind-body relation. ‘Organic’ denotes those psychiatric disorders which have a demonstrable anatomical, physiological or biochemical basis and are classified as such and ‘non-organic’ stands for those psychiatric disorders which do not have such basis and are not so classified. See, for example, Lishman (1998).

ⁱⁱ Frankfurt speaks of ‘volitions’, rather than desires, to indicate that what is at issue here is not just mere wishing, but a motivation by which the agent actually wants to be moved. Hence, those who have second-order desires, but do not want them to be effective, are also counted among wantons (Frankfurt 1988: 16). Nothing hangs on this complication in the present context.

ⁱⁱⁱ We use identification and endorsement interchangeably in this paper. By doing this, we ignore certain difficulties to do with whether identification is understood merely descriptively or also normatively. Frankfurt’s account is normally taken to involve an affirmative judgement of the desire in question (not just a mere acknowledgement of it), but he himself later denied that this is his view (Frankfurt 2002).

References

- The Mental Capacity Act. (2005).
- Adults with Incapacity (Scotland) Act (2000)
- Mental Incapacity (1995). The Law Commission. HMSO
- Re B (Adult refusal of medical treatment). [2002] EWHC 429
- Re MB (an adult: medical treatment).[1997] 2 FLR 426.
- Re C (Adult: Refusal of Treatment). [1994] 1 All ER 819.
- Re T (Adult: refusal of medical treatment). [1993] Fam 95; [1992] All ER 649.
- Re T [2004] EWHC 1279 (Fam) .
- Re W (a Minor) [1992] 3WLR 758.
- Re P (a minor) [2003] EWHC 2327.
- R v Collins ex p brady [2000] Llyd's Reports medical 355.
- Trust A, Trust B v H [2006] EWCH 1230
- R (B)v Dr SS [2006] EWCA Civ 28)
- Aleman, A., Agrawal, N., Morgan, K. D. & David, A. S. (2006) "Insight in psychosis and neuropsychological function: meta-analysis", *British Journal of Psychiatry* 189, pp. 204-212.
- American Psychiatric Association (2000) *The Diagnostic and Statistical Manual of Mental Disorders 4th Ed., Text Revision (The American Psychiatric Association)*
- Benson, P. (1987) "Freedom and Value", *Journal of Philosophy*, 84.9, pp. 465-486.
- Berg, J.W., Applebaum, P.S., and Grisso, T. (1996) "Constructing Competence: Formulating standards of legal competence to make medical decisions", *Rutgers Law Review*, 48, pp. 345-396.
- Berlin, I. (1958) "Two Concepts of Liberty", in his *Four Essays on Liberty* , Ch. III (Oxford: Oxford University Press, 1969).
- Berofsky, B. (1995) *Liberation from Self: A Theory of Personal Autonomy* (Cambridge: Cambridge University Press).

- Bratman, M. (2007) *Structures of Agency: Essays* (Cambridge: Cambridge University Press).
- Buchanan, A. (2004) "Mental Capacity, Legal Competence and Consent to Treatment", *Journal of the Royal Society of Medicine*, 97, pp. 415-420.
- Buchanan, A.E. and Brock, D.W. (1989) *Deciding for Others* (Cambridge: Cambridge University Press).
- Buller, T. (2006) "Brains, lies and psychological explanations", in: J. Illes (ed.), *Neuroethics*, [pp. 51-60/chap4], (Oxford: Oxford University Press).
- _____ (2001) "Competence and Risk-Relativity", *Bioethics*, 15, pp. 93-109.
- Bursztajn, H.J., Harding, H.P., Gutheil, T.G., and Brodsky, B.A. (1991) "Beyond Cognition: The Role of Disordered Affective States in Impairing Competence to Consent to Treatment", *Bulletin of the American Academy of Psychiatry and Law*, 19, pp. 383-388.
- Cale, G.S. (1999) "Risk-related standards of competence", *Bioethics*, 13, pp. 131-153.
- Cairns, R., Maddock, C., David, A.S., Hayward, P., Richardson, G., Szmukler, G. & Hotopf, M. (2005) "Prevalence and predictors of mental incapacity in psychiatric in-patients", *The British Journal of Psychiatry*, 187, pp. 379-385.
- Carpenter, W.T., Gold, J.M., lahti, A.C., Queern, C.A., Conley, R.R., Bartko, J. J., Kovnick, J. & Appelbaum, P.S. (2000) "Decisional capacity for informed consent in schizophrenia research", *Archives of General Psychiatry*, 57, pp. 533-538.
- Christman, J. (1991), 'Autonomy and Personal History', *Canadian Journal of Philosophy* 21.1: 1-24.
- _____ (1989), 'Introduction', in Christman (ed.) (1989), pp. 3-23.
- _____ (ed.) (1989), *The Inner Citadel: Essays on Individual Autonomy*, Oxford: Oxford University Press.
- Churchland, P. (2006) "Moral decision-making and the brain", in: J. Illes (ed.), *Neuroethics*, [pp. 3-16/chap 1], (Oxford University Press).
- Culver, C.M. and Gert, B. (1990) "The Inadequacy of Incompetence", *Millbank Quarterly*, 68, pp. 619-643.
- Culver, C.M. and Gert, B. (2004), "Competence", In: J. Radden (ed.), *The Philosophy of Psychiatry: A Companion*, pp. 258-270 (Oxford University Press).
- Cutting, J. (1997) *Principles of Psychopathology* (Oxford: Oxford Medical Publications).

- David, A.S. (1999). "On the Impossibility of Defining Delusions". *Philosophy, Psychiatry, Psychology*. 6: 17-20.
- _____ (1990) "Insight and Psychosis", *British Journal of Psychiatry*, 161, pp. 599-602.
- Dawson, J. and Szmuker, G. (2006) "Fusion of mental health and incapacity legislation", *British Journal of Psychiatry*, 188, pp. 504-509.
- Demarco, J.P. (2002) "Competence and Paternalism", *Bioethics*, 16, pp. 231-245.
- Diesfeld, K. (2003) "Insights on "Insight": The Impact of Extra-Legislative Factors on Decisions to Discharge Detained Patients" In, K. Diesfeld and I. Freckelton. *Involuntary Detention and Therapeutic Jurisprudence*. (Aldershot: Ashgate).
- Dubler, N.N. (1985) "Some Legal and Moral Issues Surrounding Informed Consent for Treatment and Research Involving the Cognitively Impaired Elderly", In: M.B. Kapp, H.E. Pies, and A.E. Doudera (eds.), *Legal and Ethical Aspects of Health Care* (Ann Arbor: Health Administration Press).
- Dunn, L.B., Palmer, B.W. & Karlawish, J.H.T. (2007) "Frontal Dysfunction and Capacity to Consent to Treatment or Research: Conceptual Considerations and Empirical Evidence", in: B.L. Miller and J.L. Cummings (eds.), *The Human Frontal Lobes: Functions and Disorders* (Guilford)
- Dworkin, G. (1988) *The Theory and Practice of Autonomy* (Cambridge: Cambridge University Press).
- _____ (1981), "The Concept of Autonomy", in R. Haller (ed.), *Science and Ethics* (Amsterdam: Rodopi Press).
- Elliot, C. (1997) "Caring About Risks: Are Severely Depressed Patients Competent to Consent to Research", *Archives of General Psychiatry*, 54, pp. 113-116.
- _____ (1991). "Competence as Accountability." *Journal of Clinical Ethics*, 2(3), pp.167-171.
- Feinberg, J. (1988) "Autonomy", in Christman (1989), Ch. 2; originally published in his *Harm to Self*, Ch. 18 (Oxford: Oxford University Press, 1986).
- Folstein, M.F., Folstein, S.E., and McHugh, P.R. (1975) "Mini-Mental State - A Practical Method for Grading the Cognitive State of Patients for the Clinician", *Psychiatric Research*, 189, pp.189-198.
- Frankfurt, H. (2002), "Reply to Gary Watson", in S. Buss & L. Overton (eds.), *Contours of Agency: Essays on the Philosophy of Harry Frankfurt*, pp. 160-164 (Cambridge, MA: MIT Press).
- _____ (1999) *Necessity, Volition, and Love* (Cambridge: Cambridge University Press).
- _____ (1988) *The Importance of What We Care About* (Cambridge: Cambridge University Press).

- _____ (1971) "Freedom of the will and the concept of the person", *Journal of Philosophy*, 68.1, pp. 5-20; reprinted in Frankfurt (1988), Ch. 2.
- Freedman, B. (1981) "Competence, Marginal and Otherwise", *International Journal of Law and Psychiatry*, 4, 53-72.
- Fulford, K.W.M. (2004) "Insight and delusion: from Jaspers to Kraepelin and back again via Austin", in: X. Amador and A. David (eds.), *Insight and Psychosis*, pp. 51-78(Oxford: Oxford University Press).
- Fuller-Torrey, E. (2008) *The Insanity Offence* (W.W. Norton).
- Garety, P.A. & Hemsley, D.R. (1997) *Delusions: Investigations into the psychology of delusional reasoning* (Maudsley Monographs: Psychology Press).
- Gert, B. (1995), "Rationality", in: R. Audi (ed.), *The Cambridge Dictionary of Philosophy*. (Cambridge: Cambridge University Press).
- Glass, K.C. (1997) "Refining Definitions and Devising Instruments: Two decades of Assessing Mental Incompetence", *International Journal of Law and Psychiatry*, 20, pp. 5-33.
- Grisso, T. and Applebaum, P.S. (1998) *Assessing Competence to Consent to Treatment*. (New York/Oxford: Oxford University Press).
- Gunn, M.J., Wong, J.G., Clare, I.C.H., and Holland, A.J. (1999) Decision-Making Capacity", *Medical Law Review*, 7, pp. 269-306.
- Gutheil, T.G. and Bursztajn, H.J. (1986) "Clinicians' Guidelines for Assessing and Presenting Subtle Forms of Patient Incompetence in Legal Settings" *American Journal of Psychiatry*, 143, pp. 1020-1023.
- Herman, B. (1993) *The Practice of Moral Judgement* (Cambridge, MA: Harvard University Press).
- Hill, Jr., T.E. (1992) *Dignity and Practical Reason in Kant's Moral Theory* (Ithaca, NY/London: Cornell University Press).
- Hotopf, M. (2005) "Assessment of mental capacity", *Clinical Medicine*, 5.6, pp. 580-84
- Illes, J.E. (ed.) (2006) *Neuroethics* (Oxford: Oxford University Press).
- Jaspers, K. (1963) *General Psychopathology* (Manchester: Manchester University Press).
- Kant, I. [1798], *Anthropologie in pagmatischer Hinischt*, reprinted in *Gesammelte Schriften*, ed. by Königlich Preußischen Akademie der Wissenschaften, Volume VII: 119-333 (Berlin, 1902ff).

- _____ [1785] *Grundlegung zur Metaphysik der Sitten*, reprinted in *Gesammelte Schriften*, ed. by Königlich Preußischen Akademie der Wissenschaften, Volume IV: 385-463 (Berlin, 1902ff).
- Kemp, R., Chua, S., McKenna, P., and David, A. (1997) "Reasoning and Delusions" *British Journal of Psychiatry*, 170, pp. 398-405.
- Komrad, M.S. (1983), "A defence of medical paternalism: maximising patients' autonomy", *Journal of Medical Ethics*, 9, pp. 38-44.
- Korsgaard, C. (1996) *Creating the Kingdom of Ends* (Cambridge: Cambridge University Press).
- Kress, K. (2004) "Why lack of insight should have a central place in mental health law". in: X. Amador and A. David (eds.), *Insight and Psychosis*, pp. 257-279 (Oxford: Oxford University Press, 2nd edition).
- Lishman, W.A.. *Organic Psychiatry: The Psychological Consequences of Cerebral Disorder*. 3rd Edition. (Oxford: Blackwell Science).
- Manson N. and O'Neill O. (2007) *Rethinking Informed Consent in Bioethics* (Cambridge University Press).
- Marson, D.C. (2001) "Loss of Competency in Alzheimer's Disease: Conceptual and psychometric approaches", *International Journal of Law and Psychiatry*, 24, pp. 267-283.
- _____ & Harrell, L. (1999a) "Executive dysfunction and loss of capacity to consent to medical treatment in patients with Alzheimer's Disease", *Seminars in Clinical Neuropsychiatry*, 4, pp. 41-49.
- _____ & Harrell, L. (1999b) "Neurocognitive changes associated with loss of capacity to consent to medical treatment in patients with Alzheimer's disease" Park, D.C. (Ed); Morrell, R.W. (Ed); *et al.*, *Processing of medical information in aging patients: Cognitive and human factors perspectives*, pp. 109-126 (Mahwah, NJ, US: Lawrence Erlbaum Associates, Publishers)
- _____, Cody, H.A., Ingram, K.K. & Harrell, L.E. (1995). "Neuropsychologic predictors of competency in Alzheimer's disease using a rational reasons legal standard", *Archives of Neurology*, 52, pp. 955-959.
- McEvoy, J.P., Applebaum, P.S., Apperson, L.J., Geller, J.L., and Freter, S. (1989) "Why Must Some Schizophrenic Patients Be Involuntarily Committed? The Role of Insight", *Comprehensive Psychiatry*, 30, pp. 13-17.
- Meyers, D.T. (1989), *Self, Society, and Personal Choice*, New York: Columbia University Press.
- Moore, M.S. (1984), *Law and Psychiatry: Rethinking the relationship*, (Cambridge: Cambridge University Press).

- Moorhead, S. and Turkington, D. (2002) "Role of emotional capacity in consent should be clarified" *British Medical Journal*, 325, p. 1039.
- Okai, D., Owen, G., McGuire, H., Singh, S., Churchill, R. & Hotopf, M. (2007). "Mental capacity in psychiatric patients: Systematic review", *British Journal of Psychiatry*, 191, pp. 291-297.
- O'Neill, O. (2003) "Autonomy: The Emperor's New Clothes", *Proceedings of the Aristotelian Society*, Supplementary Volume, 77.1, pp. 1–21.
- Owen, G.S., Cutting, J. & David, A.S. (2007) "Are people with schizophrenia more logical than healthy volunteers?", *British Journal of Psychiatry*, 191, pp. 453-454.
- _____, David, A.S., Richardson, G., Szuker, G., Hayward, P. & Hotopf M. (2008a) "Mental Capacity, Diagnosis and Insight", *Psychological Medicine*, [In Press].
- _____, R. G., David, A., Szuker, G., Hayward, P. & Hotopf M. (2008b) "Mental capacity to make decisions on treatment in people admitted to psychiatric hospitals: a cross sectional study", *BMJ (British Medical Journal)*, 337, a448.
- _____, Cutting, J. & David, A. S. (2007) "Are people with schizophrenia more logical than healthy volunteers?", *British Journal of Psychiatry*, 191, pp. 453-454.
- Palmer, B.W., Dunn, L.B., Appelbaum, P.S. & Jeste, D.V. (2004). Correlates of treatment-related decision-making capacity among middle-aged and older patients with schizophrenia", *Archives of General Psychiatry*, 61, pp. 230-236.
- Perkins, R. and Moodley, P. (1993) "The arrogance of insight?", *Psychiatric Bulletin*, 17, pp. 233-234.
- President's Commission (1982) *President's Commission for the study of Ethical Problems in Medicine and Biomedical and Behavioral Research, Making Health Care Decisions* (Washington: U.S. Government Printing Office).
- Raz, J. (1986) *The Morality of Freedom* (Oxford: Oxford University Press).
- Richardson, G. (2007) "Balancing autonomy and risk: a failure of nerve in England and Wales?", *International Journal of Law Psychiatry*, 30, pp. 71-80.
- Roth, L.H., Applebaum, P.S., Sallee, R., Reynold, C., and Huber, G. (1977) "Tests of Competency to Consent to Treatment", *American Journal of Psychiatry*, 134, pp. 910-913.
- Rudnick, A. (2002), "Depression and competence to refuse psychiatric treatment", *Journal of Medical Ethics*, 28, pp. 151-155.
- Sherlock, R. (1983) "Consent, competency and ECT: some critical suggestions", *Journal of Medical Ethics*, 9, pp. 141-143.
- Strawson, P.F. (1962) "Freedom and Resentment", *Proceedings of the British Academy*, 48, pp. 187-211.

- Stromberg, C.D. & Stone, A.A. (1983) "A model state law on civil commitment of the mentally ill", *Harvard Journal of Legislation*, 20, pp. 275-396.
- Szasz, T.S. (1999) "Medical Incapacity, Legal Incompetence and Psychiatry", *Psychiatric Bulletin*, 23, pp. 517-519.
- Tan, J. & Hope, T. (2008) "Treatment refusal in anorexia nervosa: a challenge to current concepts of capacity", in: G. Widdershoven, J. McMillan, T. Hope and L. van der Scheer (eds.), *Empirical Ethics in Psychiatry* [pp. 187-210.chap 13] (Oxford: Oxford University Press).
- Taylor, J.S. (ed.) (2005) *Personal Autonomy: New Essays on Personal Autonomy and its Role in Moral Philosophy* (Cambridge: Cambridge University Press).
- Watson, G. (1975) "Free Agency", *The Journal of Philosophy*, 72.8, pp. 205-220.
- Wicclair, M.R. (1991) "Patient Decision-Making Capacity and Risk", *Bioethics*, 5, pp. 91-104.
- Williams, B. (1973) *Problems of the Self* (Cambridge: Cambridge University Press).
- Wulff, H. (1995) "The inherent paternalism in clinical practice", *The Journal of Medicine and Philosophy*, 20, pp. 299-311.